ANALYSIS OF COMPANY-LEVEL INNOVATION PROCESS BASED ON SIMULATION OF A PERCOLATION MODEL

BY

MICHAEL J HALLORAN

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Industrial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign 2016

Urbana, Illinois

Adviser:

Associate Professor Qiong Wang

ABSTRACT

Innovation has become a key to success for many companies. By investing in research and development (R&D) activities, these companies create new products, processes, patents, and even ground-breaking theories. Innovations deliver the necessary fuel companies need to grow their revenues and profits. On the other hand, their investments are also subject to budget constraints and bear the risk of uncertain returns. It is important to understand how the innovation process will play out in various environments. This paper addresses this issue by modeling innovation as a percolation process and carrying out simulations of this model.

Our percolation model is a generalization of the one in the existing literature. The original model represents the technology space by a lattice and formulates innovation as a search process that creates a connected path in that space. We generalize this model by introducing a revenue generation function, which differentiates R&D success (formation of a connected path in the technology space) and commercial success (the paths endpoint generating a varying amount of revenue), and imposing a self-financed wealth evolution process. We conduct simulations on this generalized model and analyze the subsequent evolution of the company's wealth. We perform multiple simulation runs in different parameter regions and under different company-determined decisions. These decisions include the monetary effort (from their budget) and scope (specific area of the technological field) of the company's R&D search.

Our findings can be summarized as follows. A company's wealth over time appears to follow a "takeoff" trend similar to sales of a new product. A company's wealth initially declines, there is then a period of rapid growth, and finally wealth levels off to a steady-state value. The model also reveals that very small companies are at a disadvantage when compared to larger ones. A smaller company's limited budget and narrower scope for R&D search leads to poorer performance. Lastly, introducing some variations to the model demonstrates that companies who pool their revenue together exhibit higher performance than companies who attempt to encourage competition by only giving an innovation's revenue to the R&D group that actually discovered it.

ACKNOWLEDGEMENTS

First of all, I want to thank all of the professors, lecturers, teaching assistants, and lab assistants I have worked with at the University of Illinois at Urbana-Champaign. The remarkable undergraduate and graduate education I have received at this university has pushed me to excel not only in my academic pursuits, but life in general. Without obtaining such a strong educational foundation this work would not have been possible.

I want to give a special thanks to my adviser Professor Qiong Wang. His constant support and motivation during this project was essential for its completion. The incredible learning environment that he fosters as a teacher, along with the impeccable guidance he provides as an adviser, are perfect models I plan to follow in my future endeavors.

This work is partially funded by the National Science Foundation (NSF) grant SMA-1360188 and I thank the NSF for its support.

TABLE OF CONTENTS

1. Introduction

We model and simulate an innovation process that takes place in a profit-driven company under budget constraints. Our model accommodates the following features of an innovation process: innovations are built upon previous progress, they are random and unpredictable in nature, and they are the result of multi-disciplinary efforts. Our work is based on a generalization of percolation models developed by Silverberg & Verspagen (2003, 2005, 2007). In this section, we will start by describing the process that motivates the percolation model, followed by a review of previous development in the literature. We will then explain why percolation models provide the correct framework for analyzing the innovation process and highlight our new contributions on this front.

1.1 The Innovation Process

Innovation has been the engine of industrial progress and technological advancement. Companies undertake innovative activities to create new products, algorithms, patents, methods, and/or novel theories. Understanding the innovation process bears great value for innovating companies, but the analysis is often hindered by the fact that the nature of innovations dictate that their occurrence cannot be predicted by a formal model. Nevertheless, there is a large body of empirical research that discerns general patterns and rules of the innovation process based on: 1.) Historical data (e.g. Trajtenberg 1990 using patent records) 2.) Case studies. The consensus views are:

- **Innovations are Clustered in Time**
  - Many time periods will exhibit little or no innovations and then suddenly many innovations will occur in a single time period.
  - "[Innovations] are not evenly distributed in time, but … on the contrary they tend to cluster, to come about in bunches, simply because first some, and then most firms follow in the wake of successful innovation" (Schumpeter 1939, p. 75)
- **Innovations are Clustered in the Technology Space**[1]
  - When innovations suddenly appear they are normally from the same or closely related technological field.
  - Thought to be due to the localized nature of research & development and innovations following the specific needs of consumers at the time.

[1] See Foray and Grubler (1990) ; Saviotti (1996) ; Frenken and Leydesdorf (2000)

- **Innovation Size is Highly Skewed**
    - Most innovations are minor (only slightly better than current technology), but some are major (much better and more complex than current technology).
- **Revenue Generated by Innovations is Highly Skewed**[2]
    - Most innovations produce little or no revenue. The majority of revenue comes from a select few innovations.
- **Innovations Build from Previous Innovations** (Not a Totally Random Process)
    - Innovations are incremental and require past ideas, processes, and parts to function.
    - Innovative products become more attractive to develop when the initial outcome has gained a foothold in the marketplace. The market diffusion of the innovation becomes stronger with more and more product improvements (because consumers want a better product). Thus a cycle of consistent innovative improvement (incremental innovations) is created (Silverberg and Verspagen 2003).

In this paper we wish to use the innovation process as a test bed for experimentation. It is therefore necessary to go beyond empirical data and build a model that provides the flexibility for experimentation while also accurately capturing the aspects of the innovation process discussed above. The next section gives an overview of a class of models using percolation theory that researchers have been using to simulate the innovation process.

1.2 Summary of Silverberg & Verspagen's Percolation Models

The models discussed in this paper are based upon percolation theory, which is the study of how clusters connect within a given random space. Percolation is ubiquitously observed in nature itself, with a specific example occurring on rainy days. A concrete surface (i.e. random space), is initially dry until water droplets (i.e. clusters) begin to fall onto the surface. As the rain begins to fall, initially there are just a few individual wet dots visible on the concrete. Eventually some of the dots begin to connect together to form groups of wet patches (i.e. chains of clusters). This act of clusters connecting to form chains is the key concept regarding percolation models.

An early application of percolation theory to technological change can be found in Cohendet and Zuscovitch (1982). Gerald Silverberg and Bart Verspagen also developed models to simulate the innovation process (2003, 2005, and 2007). Below we give a brief review of these models.

[2] See Scherer (1998) ; Harhoff et al. (1999) ; Scherer et al. (2000) ; Harhoff et al. (2003)

Within those models the random space is represented by a rectangle called the "lattice." The clusters of the model are the individual sites that make up the lattice itself. These sites can simply be viewed as smaller rectangles making the model look like a checkerboard (Figure 1).

The lattice is bounded in the horizontal axis with a finite number of columns (n) and unbounded in the vertical axis. This means that each individual lattice site is uniquely identified by its row (i) and column (j) numbers.

The horizontal axis represents the technology space itself with each column signifying a unique technological field (computer science, chemistry, etc.). Similar fields are grouped more closely together. The vertical axis represents performance with the higher rows representing more complex technologies.

Along with the coordinates (i, j), each lattice site also has a state. These states are the critical element that allows the percolation model to be representative of the innovation process. In Silverberg and Verspagen's models there are at most four states that a lattice site can be in.

0. Not Discoverable (excluded by nature)
1. Discoverable
2. Discovered (but not yet viable)
3. Viable

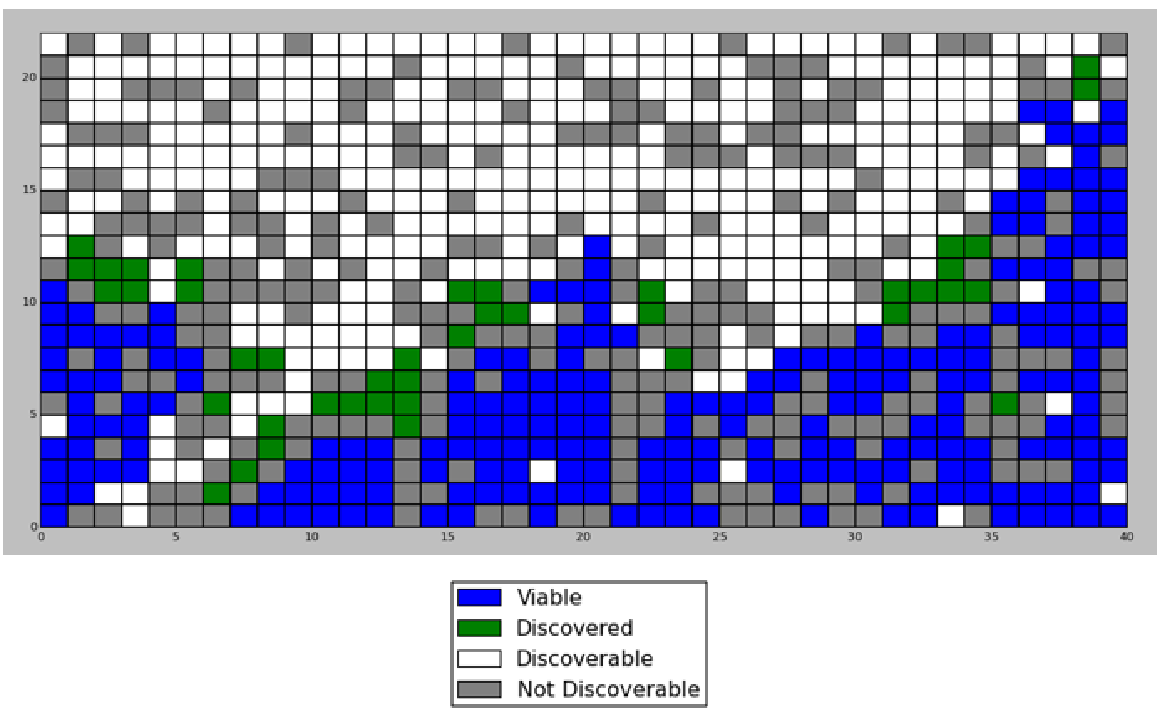


Figure 1: Lattice Example

The first two states are properties of the technological space itself and are unknown to the agents of the model in the beginning. If a state is excluded by nature then it can never become an innovation. The site will always remain in this state of exclusion. On the other hand, if a site is discoverable (i.e. possible) then it can eventually become discovered and finally viable. This last type of state, viable, is equivalent to a site becoming an "innovation." The "best practice frontier" (BPF) of a column is the most advanced innovation that technological field currently has to offer. This is the row value of the highest "viable" site in a given column.

Research and development (R&D) search is how agents of the model gain information about the existing technology space around them. In Silverberg and Verspagen's models, R&D search is a step-by-step process that starts at a column's existing BPF site. Given a search radius (r), the search is conducted at sites in a diamond pattern around a column's BPF. This pattern and how it changes with the search radius can be seen in Figure 2.

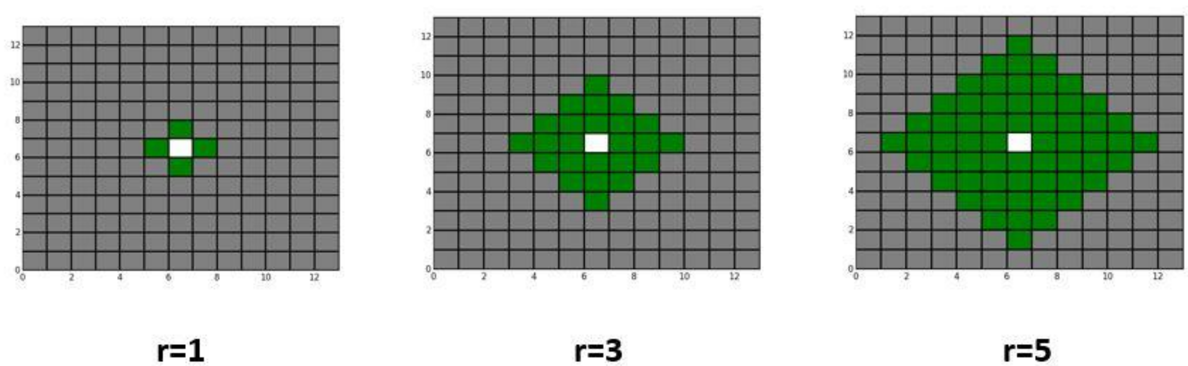


Figure 2: R&D Search Pattern (White Site = BPF)

Research & development is how a lattice site moves from state possible (1) to state discovered (2). With some probability, if R&D search is conducted on a site currently in the possible state (1) then it will change to state discovered (2). This probability is dependent upon the effort or capital that agents have placed into R&D that period.

The last state change to consider is from state discovered to state viable. This state change represents when a theory, idea, concept, etc. becomes useful to a company. This state change takes place if and only if the site has a "neighbor" already in state 3. A lattice site has at most four neighbors (top, right, left, and bottom). In this way, the previously state 2 site becomes part of a connected chain of state 3 sites. When this state change from 2 to 3 happens an innovation is formed. The size of the innovation depends on how big of a jump from the current BPF the new BPF will be.

1.3 How the Percolation Model Captures Aspects of the Innovation Process

Now that the framework of Silverberg and Verspagen's percolation models has been discussed, the question remains: Does this model accurately capture the known aspects of the innovation process described in Section 1.1? Below we discuss how each of aspect is reflected by the formulation of the percolation model.

- **Innovation Build from Previous Innovations**
    - Given by the constraint that in order for a lattice site to become viable it must have neighboring sites that are also viable.
    - This creates a chain of innovations building upon one another from all the way at the bottom of the lattice structure (the most primitive of innovations).
- **Innovation Size is Highly Skewed**
    - There are many instances (frequency) of small innovations and very few instances of large innovations (Figure 3).

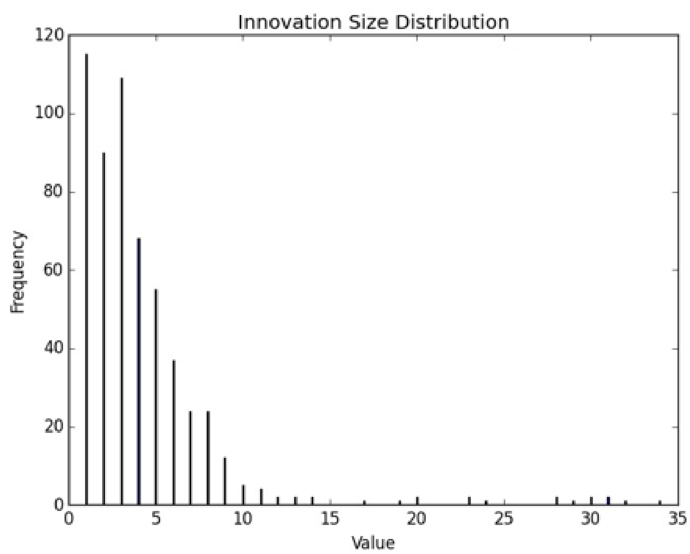


Figure 3: Histogram of Innovation Size[3]

- **Innovations are Clustered in Time**
    - Figure 4 illustrates how a large number of innovations can be found at once in the model. It happens when a large cluster of state discovered sites finally become connected to the technological frontier (BPF) due to a "cornerstone innovation." This turns all of them to state viable instantly.

[3] Figure from a single simulation run that replicates the Silverberg and Verspagen (2005) model

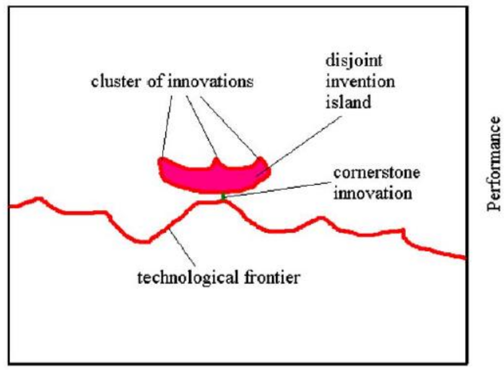


Figure 4: Lattice Depicting a Cluster of Innovations Being Formed[4]

- **Innovations are Clustered in the Technology Space**
    - The R&D search radius limits the distance from the BPF a new innovation can be found. This corresponds to limiting how different a new innovation can be from a previously proven one.
- **Innovations are Unknowable Before they Happen**
    - Initial states are unknown to agents of the model until discovered by R&D.

In our work, we extend upon the models of Silverberg and Verspagen by taking a company-specific view of the innovation process. This work takes into account that not every discovered innovation will result in a financial windfall for the company. We introduce the new concept of a lattice site being a "revenue generator", and thus a commercial success. Further extensions include the consideration that a company's R&D search is constrained by budget dynamics and performance can be analyzed by looking at a company's wealth (not merely the BPF). Under these extensions we carry out simulations to evaluate performance of the innovating company.

The rest of the paper is organized as follows. In Section 2 we discuss the framework of our model, how the model conducts research and development, we introduce key metrics, and compare/contrast our model with those of Silverberg and Verspagen. Section 3 is devoted to explaining how simulation runs of our model were conducted and provides an overview of the algorithm used. Section 4 provides the results of our simulation runs and gives insights into the meaning behind them. Section 5 outlines some variations to our model and the effects these changes had on simulation results. We give a brief summary of the paper, draw conclusions, and offer possibilities for further research in Section 6.

[4] Figure comes from Silverberg and Verspagen (2005)

2. A Model of Innovation Process

In this section, we present our model of the innovation process. We envisioned this model to imitate the decision making a company makes in their search for innovations. The company has the opportunity to perform research & development in an array of technical fields. The question the company wants to answer is: How should we divide our R&D effort amongst our technological field frontier to maximize performance? The effort to search for these revenue generators is subject to budget constraints. The company is rewarded by discovering revenue generators (patents, physical products, etc.) that are "found" by investing some capital into R&D search throughout the technological frontier. The company doesn't know which of their research endeavors will be successful, and if successful, how much revenue a discovery will produce.

Consequently, our model is composed of three major components which will be elaborated upon in the following subsections:

1. A structural layout of the innovation space
2. A search process applied to the space
3. A measure of financial performance as a result of the search

2.1 Structural Layout

Like the models discussed in Section 1, our framework creates a lattice structure within which the innovation process can be replicated by continuous percolation. Our lattice has a finite number of columns (n), each representing a specific technological field. The lattice is unbounded in the vertical direction (i.e. number of rows), with higher rows (i.e. higher number) representing more advanced innovations. The "height" of the lattice is unbounded so as not to impose an artificial constraint on how advanced an innovation can be.

A lattice point can be described by giving the row and column where it resides (i, j). Points within the lattice structure are referred to as "sites." There are three types of sites:

1. Not Feasible ($S_{i,j} = 0$)
    a. Innovation not possible (excluded by nature)
2. Feasible (Non-Revenue Generator) ($S_{i,j} = 1$)
    a. Innovation possible, but produces no monetary return
3. Feasible (Revenue Generator) ($S_{i,j} = 2$)
    a. Innovation possible, and produces a monetary return

A set of probabilities determines how many of each type of site is expected within the lattice. The percolation probability (q) dictates if a site will be Feasible ($S_{i,j} = 1$; $S_{i,j} = 2$). This means that (1-q) is the probability that a site is type Not Feasible. This probability q comes from percolation theory itself and will be discussed further in later sections. In order to determine between the two types of feasible sites a "revenue generator" probability (p) is used. Feasible sites are revenue generators with probability p. This seeding of the lattice with revenue generators creates a kind of "hide and seek" game. The revenue generators are hiding somewhere within the lattice and it is the goal of the company to find as many of them as possible. This is meant to capture the reality that only a select few innovations (i.e. viable sites) will actually give the company a return on their investment.

Site Type Probabilities:

- Not Feasible
    - $(1-q)$
- Feasible (Non-Revenue Generator)
    - $q(1-p)$
- Feasible (Revenue Generator)
    - qp

Site Attributes:

1. Resistance ($R_{i,j}$)[5]
    - This is a value R&D effort must overcome for the company to discover the site
    - The value follows a distribution with mean R_{μ} and standard deviation R_{σ}
    - Unknown to the company
2. State Perceived by the Company ($L_{i,j}$)
    - **Hidden** ($L_{i,j}$ = -1)
        - Unknown to the company
        - Can be any type of site
    - **Not Feasible** ($L_{i,j}$ = 0)
    - **Feasible** ($L_{i,j}$ = 1)
        - Known to the company and useful
        - Can be Feasible (Non-Revenue Generator) or Feasible (Revenue Generator)
    - **Viable** ($L_{i,j}$ = 2)
        - Innovation

[5] Concept introduced in Silverberg and Verspagen (2007)

- Company now knows if the site generates revenue or not

Note that these states are slightly different than those discussed previously in Section 1.

3. Revenue ($z_{i,j}$)
   - Value the company earns when the site becomes viable ($L_{i,j} = 2$) to the company:
     - Not Feasible – Not Applicable
     - Feasible (Non-Revenue Generator) - Zero
     - Feasible (Revenue Generator) - distribution with mean p_μ and standard deviation p_σ
   - Unknown to the company until site's state becomes viable

It is important to see the distinction between a site's type and its state. The type is what the lattice site actually represents, while the state is how the company currently perceives it. The company is totally unaware of the site's type while it is in the initial "Hidden" state. Thus hidden sites are the only sites where the R&D budget is spent on because these are the sites with the unknown information. Once a hidden site has enough research cost put into it to overcome the associated resistance value, the state either changes to "Not Feasible" or "Feasible." Sites that become "Feasible" may eventually become "Viable" if they become connected to a branch of sites that are already "Viable." Once the state "Viable" is reached, the company knows whether or not the feasible site generates any revenue.

Thus the possible state dynamics that occur when a site is discovered (i.e. the resistance is overcome) during R&D search are as follows:

State Hidden (-1):

- To state Not Feasible (0)
- To state Feasible (1)

State Not Feasible (0):

- No further state changes are possible

State Feasible (1):

- To state Viable (2) if and only if an element of $L_{neightbor} = 2$ where $L_{neighbor} = \{L_{i+1,j} ; L_{i-1,j} ; L_{i,j+1} ; L_{i,j-1}\}$

2.2 The Research & Development Search Process

This section describes how our model attempts to capture how a company's research and development affects the technological landscape. This is the key process of the model in that it is the driving factor behind why state changes occur. The goal of research and development is twofold:

1. **Exploration**
    - Discover the feasibility (or lack thereof) of sites currently hidden ($L_{i,j}$ = -1) from the company
    - This discovery occurs when the resistance attribute of a site becomes negative ($R_{i,j}$ < 0)
2. **Exploitation**
    - Turn feasible sites ($S_{i,j}$ = 1 or 2) into state viable ($L_{i,j}$ = 2) to collect (probabilistically) revenue

The R&D search process is conducted at every column (i.e. n times in total) each time period (t). Each column conducts their search around a central lattice site. To determine what lattice site to use as the center we introduce the concept of a column's best practice frontier. A column's best practice frontier (BPF_j) is introduced in Silverberg and Verspagen (2005) and defines that column's most complex innovation. The BPF can be represented as an array of size n with values following:

$$BPF_j = \max_i (i \mid L_{ij} = 2) \quad for \quad j \in [0, n) \tag{1}$$

In other words, the BPF of a column is the row value of the highest state 2 lattice site. If no site of state 2 (viable) exists in a given column then the BPF is set to -1. The BPF is important because it will serve as the point around which R&D search is conducted. If the BPF is equal to -1 then that column does not perform R&D search. In order to have an initial BPF array for R&D search to occur, a portion of the feasible sites in the bottom row are set to state viable. This gives initial lattice points that R&D search can be conducted around.

The search is conducted with a search radius (r) branching off from the central BPF site. This creates a diamond shaped area over which R&D search is conducted (see Figure 2).

The number of sites (num_{sites}) within the search area falls within the range [2r, 2r(r+1)]. Figures 5 & 6, where the white site represents the BPF, demonstrate two other properties of the R&D search area:

- Search Areas Can be Truncated by the Lattice Boundaries
- Search Areas Can Overlap

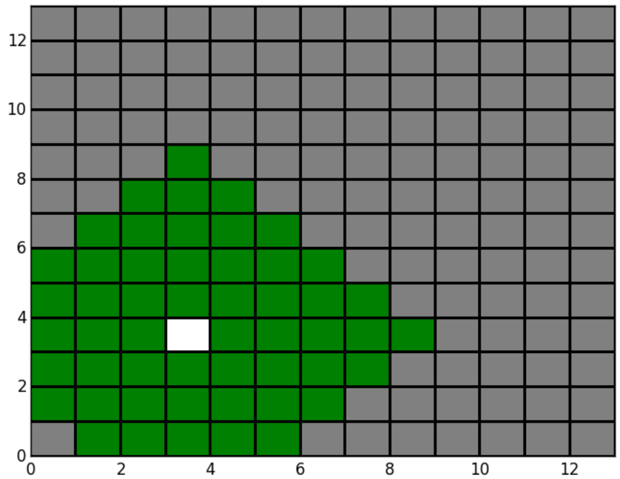


Figure 5: Truncated Search Area Example

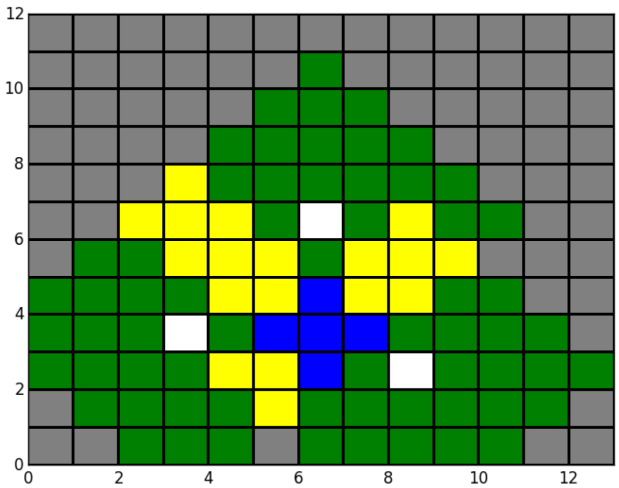


Figure 6: R&D Overlap (Green Searched Once; Yellow Searched Twice; Blue Searched 3 Times)

In order for R&D to operate it must draw from a budget, that the company initially sets, devoted for research and development purposes. For simplicity, this total budget (B) for R&D is then divided equally amongst all fields (columns) for the R&D search according to (2). In reality the company could allocate their budget many different ways amongst the columns. The company only directly gives capital for R&D once. After this, the search process must rely on the remaining budget and additional funding from discovering revenue generators. In this regard the columns are self-financed by their innovation discoveries.

$$b = \frac{B}{n} \tag{2}$$

Once the budget is allocated to every column, a decision must be made on how much of their individual budget (b) a column should spend this period. In this model, that decision made by each column is simply a percentage ($b_{\%}$) with every column making the same percentage decision. This decides the actual monetary effort (E) that a column puts into R&D each period subject to some constraints.

$$E = min(max(b * b_{\%}, E_{min}), b) \tag{3}$$

For simplicity, this effort (E) is then equally distributed amongst all hidden lattice sites (num_{-1}) within the search radius. This is unique to this model. In previous models the effort was simply determined by the number of sites in the search radius no matter if their feasibility had already been discovered. In this model research & development search is only concerned with sites with $L_{i,j} = -1$.

Thus the actual R&D monetary effort that each individual site receives from a given column's search is given by (4). This is the key value that will decide whether or not the hidden lattice site will become discovered and change states. This value is what is deducted from each site's resistance value ($R_{i,j}$). Therefore, the larger the value of E_{site} the more likely the site will become discovered by the company.

$$E_{site} = \frac{E}{num_{-1}} \tag{4}$$

To determine whether the lattice site should change from its hidden (-1) state, the resistance value of each site in the search radius is decreased according to (5). Thus if the resistance is a positive value, and E_{site} is of greater value, the site will no longer be hidden.

$$R_{i,j} = R_{i,j} - E_{site} \tag{5}$$

If the resistance value becomes negative then the state of lattice site becomes either Feasible ($L_{i,j} = 1$) or Not Feasible ($L_{i,j} = 0$). It is important to note that a given sites resistance value can be decreased multiple times in a given time period due to overlapping of R&D search of different columns.

The progression of state changes in a period can lead to the discovery of one or more revenue generators. Revenue generators produce immediate capital that the company adds back to their total budget for R&D (B). The total revenue generated during a single column's R&D search (z_j) is the sum of all the revenue values ($z_{i,j}$) of sites that changed from state 1 to state 2 during the R&D search process. In the real world not all of the revenue generated by these sites would go immediately back to R&D, but consider that this model tailors the parameters so that the revenue accurately captures how much money the company would put back into the total R&D budget.

In this way, the total R&D budget (B) and individual budgets (b) are constantly changing from time period to time period and are dependent on these parameters.

1. Current total budget (B_t)
2. Percent of the total budget to use this period ($b_{\%}$)
    a. Set initially and fixed
3. Number of "hidden" sites located in the diamond search area (num_{-1})
4. Minimum effort per column per period (E_{min})[6]
    a. Set initially and fixed

$$B_{t+1} = B_t - (E_t * n) + \sum_1^n z_j \qquad (6)^7$$

$$b_t = \frac{B_t}{n} \qquad (7)$$

$$E_t = min(max(b_t * b_{\%}, E_{min}), b_t) \qquad (8)$$

Note that r, $b_{\%}$, and E_{min} are constant across all time periods.

The company will continue to perform R&D search until one of the following conditions are met:

1. The total R&D budget (B) becomes zero.
2. The number of periods (t) exceeds the maximum number set by the user (t_{max}).

Now that the R&D search process has been described, it is apparent that there are multiple decisions a company must make in conducting this search. In this way, different strategies can be formed to tailor to the specific needs of a given company. The two main strategic decisions a company has to make are:

1. Search Radius (r)
2. Budget Allocation ($b_{\%}$)

A larger search radius means that R&D covers more space and the potential for search areas to overlap is greater. However it also means that a column's effort (E) is spread amongst a larger number of sites making the effort per site (E_{site}) smaller. A larger $b_{\%}$ means that more of a

[6] This minimum effort is necessary to allow for the total budget to reach zero.

[7] Note in (6) that the effort is multiplied by the number of columns (n). This means as soon as the effort is allocated to a column it is lost. This holds even if the BPF of the column is -1 or there are no sites hidden from the company within that column's search area.

column's budget (b) will be put into the R&D search effort (E) in a given time period. However, a greater $b_{\%}$ value could lead to the company quickly losing its entire budget.

The natural question to ask is: What values for r & $b_{\%}$ should the company choose for them to maximize their performance? This is a complicated question which Section 4.2 will discuss in much greater detail. First, we must introduce the metrics by which the company's performance will be assessed.

2.3 Defining Performance Metrics

In previous percolation papers by Silverberg and Verspagen (2005, 2007) some key metrics to analyze were the maximum height achieved within the lattice, the number of innovations that occurred, and the size of these innovations. This paper introduces two new metrics which focus on the fact that this model is looking at the innovation process through the eyes of a company.

1. Bankruptcy (bank)
2. Wealth (W_t)

The metric "Bankruptcy" will denote simulation runs where the company at some period t before t_{max} depleted all of their R&D budget. It is simply a binary variable with zero (one) meaning it didn't (did) occur. In the real world there may be possible ways to continue when this occurs (e.g. receiving a loan), but for the purposes of this model if this condition is met all research and development activity ceases immediately. One cause of bankruptcy can be connected to the concept of "deadlock" from Silverberg and Verspagen (2005). This occurs when the best practice frontier of the lattice is completely enclosed by sites not feasible for innovation to occur. This concept of "deadlock" is important because it means that no matter how much monetary effort is put into R&D, or how large the search radius is, no further innovations will be discovered (Figure 7).

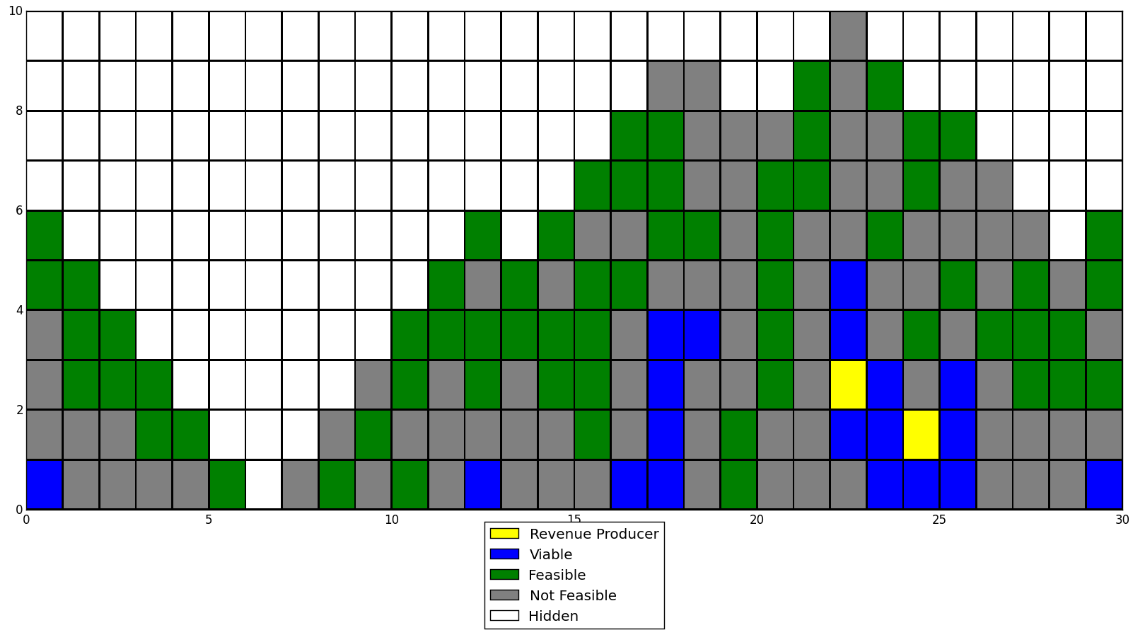


Figure 7: Example of "Deadlock"

"Deadlock" is closely associated with the percolation probability (q) and is merely one reason that bankruptcy may occur. "Deadlock" and the other causes of bankruptcy will be discussed in greater detail in Section 4.

A company's wealth (W_t) is an important metric, especially when bankruptcy does not occur. In reality, a monetary measure is how most companies measure their performance. The wealth metric is calculated during each period. This allows for tractability while performing an actual simulation run. It is represented as a proportion of the current total budget to the initial total budget according to (9).

$$W_t = \frac{B_t}{B_0} \tag{9}$$

This proportion was constructed to be a simple measure of the "profitability[8]" of the company. If wealth is greater than one, it means the company has generated more revenue than it has spent on R&D. Obviously when bankruptcy does occur this metric will be equal to zero. Section 4 will use these metrics to analyze the simulation runs and explain trends the data exhibits.

[8] This model looks at a company solely through the scope of research and development and the budget associated with the process. This profitability does not reflect the actual amount of money a company will earn in reality, but is rather meant to show data trends.

2.4 Similarities & Differences with Silverberg & Verspagen's Models

Our model combines concepts from all three previous works by Silverberg and Verspagen (2003, 2005, and 2007) and also introduces new concepts to fit our specific needs/goals. This section is devoted to listing the key common features held by both our models as well as the key features that are unique to this model.

Common Features

- Use of percolation theory to represent the innovation process
- Representation of the technological frontier as a lattice structure with a finite number of columns and unbounded number of rows
- Site Types: Not Feasible & Feasible
- Site States: Not Feasible, Feasible, and Viable
- Site's resistance attribute definition
- Best Practice Frontier (BPF) definition
- Search pattern (diamond shape with radius r)
- Requirement that a neighboring site (up, down, left, right) be viable for a site to become viable as well
- Maximum height metric

Our Model's Unique Features

- Model's viewpoint of the lattice as a single company capable of R&D across a number of technological fields (columns).
- Site Type: Feasible (Revenue Generator)
    - Distribution of these sites throughout the lattice with probability p
- New site state: "Hidden"
    - Important so that R&D search only focuses on these sites unlike Silverberg & Verspagen's models where R&D focused on all sites within the search area.
- R&D search only concerning sites still "hidden" from the company
- Dynamically changing R&D budget [9]
    - Concept that innovations produce revenue that goes back into a R&D budget
- Bankruptcy & Wealth Metrics

[9] A concept for R&D effort being dependent on finding innovations is presented in Silverberg & Verspagen (2007), but is very different from the one presented in this paper.

3. Simulation

This section goes into the specifics of how computer simulations of our model were conducted. It is meant to provide specifics such that readers can reproduce the simulations if they desire. To program the model, a script was written using Python programming language. The lattice structure of our percolation model is represented by matrices. Elements within these matrices represent the sites and their corresponding attributes. The user inputs values for the simulation parameters (Table 1). This allows each of these parameters to be varied for the user to experiment with different combinations of values. The script then runs and outputs statistics based upon the metrics discussed in the previous section. Multiple runs can be performed once the user has set their desired parameter values (num_{runs}). In this model, the values of the metrics are averaged over all runs and then given as an output.

Table 1: List of User Controlled Inputs

Simulation Inputs	
▪ Percolation Probability (q)	▪ Initial Budget per Column ($b_{initial}$)
▪ Number of Columns (n)	▪ Budget Percent for R&D ($b_{\%}$)
▪ Search Radius (r)	▪ Minimum Effort per Column (E_{min})
▪ Resistance Distribution (R_{μ} , R_{σ})	▪ Number of Time Periods (t_{max})
▪ Revenue Distribution (p_{μ} , p_{σ})	▪ Number of Simulation Runs (num_{runs})

3.1 The Algorithm

In this section we present the simulation algorithms used to implement our percolation model. Flowcharts of the algorithm can be seen in Appendices A & B.

Step 1.) Matrix Initializations

- Creates three matrices with r+1 rows and n columns
  - S – Matrix of the actual site values (currently unknown to the company)
    1. Initialize elements: $S_{i,j} = 0$ or 1 based on percolation probability (q)
    2. If $S_{i,j} = 1$ and $i > 0$: $S_{i,j} = 2$ with probability p
  - R – Matrix of the resistance values
    1. Initialize elements: $R_{i,j} \sim \text{lognormal}(R\mu , R\sigma)$
  - L – Matrix of the state values from the company's perspective
    1. Initialize elements: $L_{i,j} = -1$
    2. If $S_{0,j} = 1$: $L_{0,j} = 2$ with probability 0.5
- Initialize the values of the BPF array

- If $L_{0,j} = 2 : BPF_j = 0$
  - Else : $BPF_j = -1$
- Initialize the value of Total Budget (B)
  - $B_0 = b_{initial} * n$
- Initialize the binary bankruptcy variable (0 – No ; 1 – Yes)
  - bank = 0

Step 2.) R&D Search *(See Appendix C for Illustrative Example)*

- Obtain the matrix coordinates within the search diamond of radius r around the central BPF.
- Determine the R&D effort per column (E) based upon (7) & (8).
- Determine the R&D effort per lattice site (E_{site}) according to (4).
- Update according to (5).
  - If $R_{i,j} < 0$: $L_{i,j} = S_{i,j}$
  - For each $L_{i,j} = 1$: Check Neighbors $_{i,j}$= $[L_{i+1,j} ; L_{i-1,j} ; L_{i,j+1} ; L_{i,j-1}]$
    - If any element of Neighbors $_{i,j}$ = 2: $L_{i,j} = 2$
      - If $S_{i,j} = 2$: Revenue generator is found and coordinates (i,j) are recorded.
      - **Two Check Function (row, column)**: row and column are function inputs
        - If any element ($x_{a,b}$) of Neighbors $_{i,j}$ = 1: $L_{a,b} = 2$
          - If $S_{a,b}$ = 2: Revenue generator is found and coordinates (a,b) are recorded.
          - For each such $x_{a,b}$: Two Check Function (a,b)
            - This creates a loop that stops when the entire branch of connected $L_{i,j} = 1$ elements have been changed to $L_{i,j} = 2$.

Step 3.) Repeat Step 2 for Each Column (Column Order is Randomized)

Step 4.) Update the BPF According to (1)

Step 5.) Update the Total Budget (B) According to (6)

Step 6.) Check for Bankruptcy

- If B = 0:
  - bank=1
  - Go to Step 9

Step 7.) Add Rows to Matrices

To help with the simulation run time the matrices only have as many rows necessary so as not to inhibit the R&D search process. This step adds rows to the matrices if in the next period the search would be inhibited (i.e. the search diamond would be truncated by the top of the lattice). This can be seen in the conditions listed below:

- If (m < max(BPF) + r): where m = matrix height
  - Add rows until m = max(BPF) + r
    - Done for all matrices (S,R,L)
    - Specific values of the rows follow the conditions set in Step 1

Step 8.) Repeat Steps 2-7 for t_{max} Periods (or Until "Bankruptcy" Occurs)

Step 9.) Repeat Steps 1-8 num_{runs} Times

Step 10.) Return Averaged Metrics to the User

- Maximum Height Achieved:

$$h_{max} = \max(BPF) \tag{10}$$

- Wealth according to (9)
- Bankruptcy Percentage:

$$bank_{\%} = \frac{\sum_{i=1}^{num_{runs}} bank_i}{num_{runs}} \tag{11}$$

4. Results & Discussions

The purpose of running these simulations is to evaluate a company's innovation performance and how various factors affect this performance. This purpose can be narrowed down into three more specific categories:

1. Examination of the external limits for the survival of the company
2. Examination of the influences different operating strategies have on performance
3. Examination of notable trends exhibited by the simulation results

Simulations were run varying either one or two variables at a time. The results are averaged across five simulation runs (i.e. $num_{runs} = 5$). The results are displayed as either two or three-dimensional plots with the height being user defined metrics (see Step 10 of the algorithm for list of metrics). This was done to better visualize and analyze any trends embedded within the simulation results. Specific values for the input parameters can be found in the boxes next to their corresponding figures.

4.1 External Limits on the Survivability of the Company

In this section, we discuss the influence of various external factors on the company's survivability. Specifically we explore a range of input values looking for limits beyond which the company is likely to experience no bankruptcy (i.e. survives). Therefore, the key metric used in this analysis is $bank_{\%}$. As defined in (11), this metric gives the percent of simulation runs where the company goes bankrupt within the t_{max} time periods.

The simulation inputs we studied can be separated into two categories:

1) **Constraints of technology space**
   a) Percolation probability (q)
   b) Revenue generator probability (p)
   c) Mean of the resistance matrix (R_{μ})
   d) Mean of the revenue generators (p_{μ})
2) **Constraints of the company**
   a) Initial budget ($b_{initial}$)
   b) Breadth of Technical Scope represented by the number of columns (n)

We consider the above factors to be "external" because they are outside of the company's decision making process. The next two subsections give insight into the values these parameters need to have in order for the company to survive.

4.1.1 Effect of Structural Constraints on Survivability

First we discuss the effect the two probabilities p and q have on the company's chance to survive. Figure 8 is a three-dimensional plot showing p & q on the bottom axes and $bank_{\%}$ on the vertical axis.

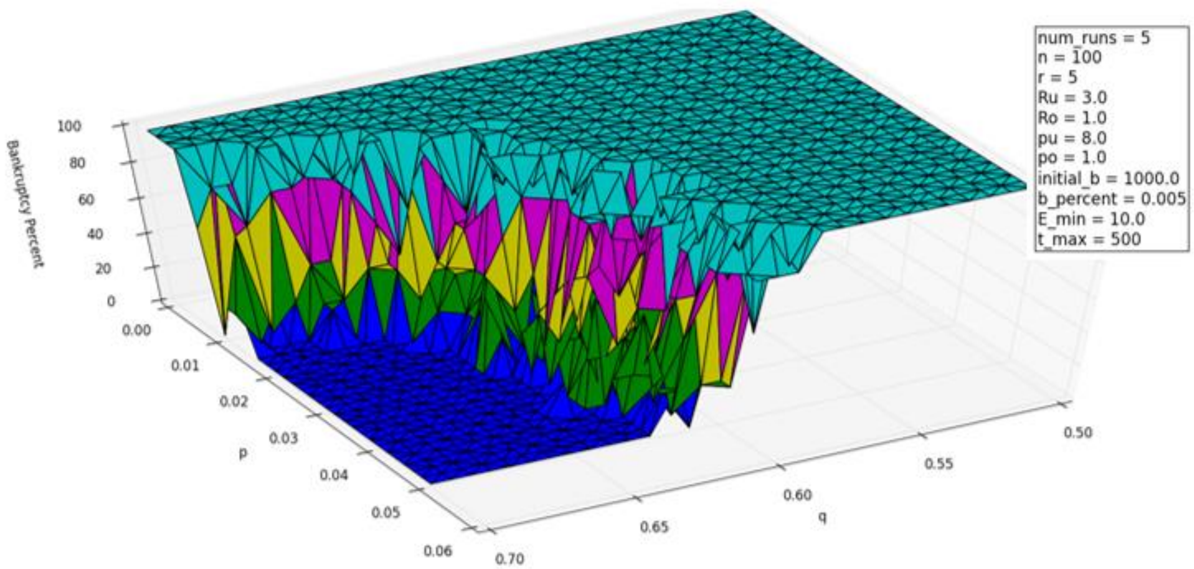


Figure 8: Bankruptcy Percent Varying Percolation Probability (q) with Revenue Generator Probability (p)

There appears to be a region of small values (p,q) for which the company will go bankrupt with near certainty. This region of values defines the external limit of survivability for the company.

If q is too low then there a large number of lattice sites of type Not Feasible. This creates blockages in the lattice that make it difficult for innovation pathways to form. This demonstrates that a "deadlock" (as discussed in Section 2.3) is occurring. No matter how much money a company puts into research & development or how large they make their serach radius, the $S_{i,j} = 2$ lattice sites cannot go around the lattice impasses ($S_{i,j} = 0$). The critical threshold (external limit) value of the percolation probability (q) is well known in percolation probability. The critical probability's (q_c) value varies depending on the type of percolation being used and graph structure. However, in Grimmett (1989) q_c for percolation models following our structure was mathmatically proven to be around 0.593. In Figure 7 the critical value for q appears to be about 0.6, which is very close to the percolation theory's proven value. This lends support that our computer simulation and model framework is accurately representing a true percolation model.

A very low p means that there are simply not enough opportunities to find revenue generators within the lattice. This leads to an insufficient amount of revenue generation that cannot recover the search expenses, thus leading to the complete depletion of the budget.

The two other constraints of the technology space are the expected values of resistance and revenue (R_μ & p_μ). Figure 9 below demonstrates the trends from varying these parameters.

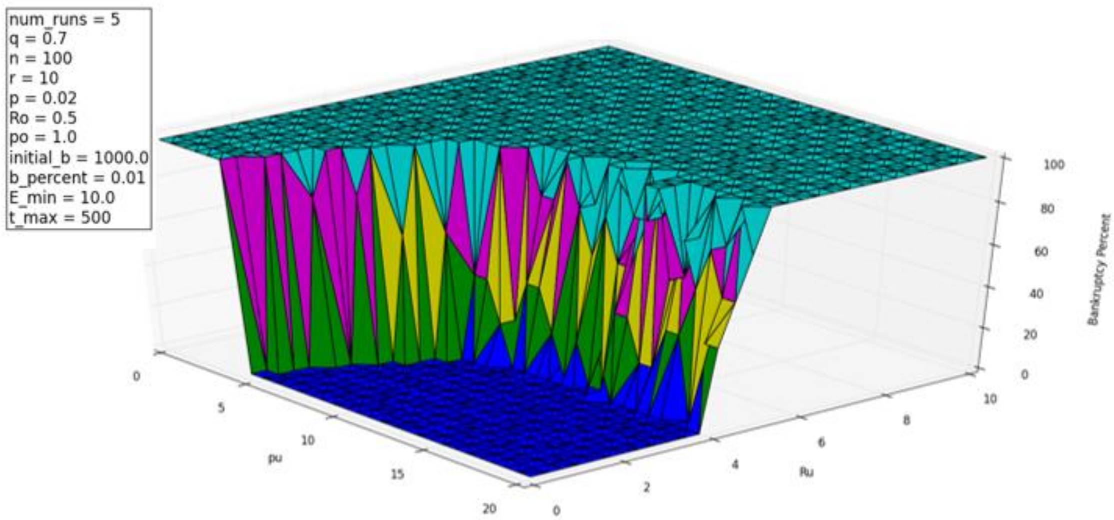


Figure 9: Bankruptcy Percent Varying Revenue Mean & Resistance Mean

There once again appears to be external limits of survivability below which a company cannot expect to stay solvent for tmax = 500 periods. This region contains small values of pμ and large values of R_μ.

Too small of revenue generators (p_μ) means that while there may not be a shortage of revenue generators, the average amount each one produces may not be sufficient to recover the search cost. On the other hand, too large of resistance from the lattice (R_μ) makes discovery of a hidden site costly. In other words, an unsustainable amount of investment is necessary before any revenue generators can be discovered.

4.1.2 Effect of Company Constraints on Survivability

Moving on from constraints due to the technology space itself, bankruptcy can also be caused by constraints of the company.

Figure 10 shows the bankruptcy percent for a range of values of the initial budget per column ($b_{initial}$) and the technical breadth of the company (i.e. number of columns (n)).

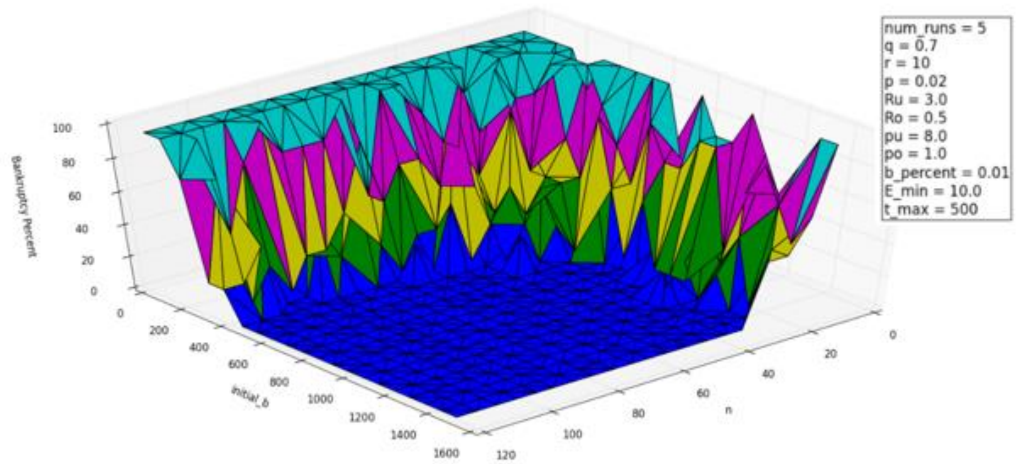


Figure 10: Bankruptcy Percent Varying the Initial Budget & Number of Columns

Once again, the figure demonstrates that there is a region of values ($b_{initial}$, n) where the company will almost certainly go bankrupt within 500 periods. This region is composed of small values of both the initial budget and technical breadth.

If the company starts from a very small budget they simply do not have the initial capital required for survival. The effort (E) that they put into research and development is too low for the discovery of revenue generators and their budget quickly becomes totally depleted. A low value of n (technical breadth) means that a company cannot performs research and development in a narrower set of technical fields. It also corresponds to the company having a smaller initial total budget (B_o) according to $B_0 = b_{initial} * n$. This smaller budget and narrower innovation search provides less flexibility for the company. This leads to the company succumbing to bankruptcy at a higher percent than those with a wider breadth of technical scope.

4.2 How Different R&D Strategies Affect Performance

This section focuses on the effects that different R&D strategies have on a company's performance. The plots in this section use the maximum height and wealth metrics as their vertical axes. As stated previously in Section 2.2, the two main input decisions that the company can strategically make regarding R&D search are:

1) Search Radius (r)
2) Budget Allocation (b%)

Figures 11-13 are the results from the same simulation run varying the input parameters r & $b_{\%}$. The other fixed parameter values can be seen in the box to the right of Figure 11 and are the same values for the following two figures. Also, when looking at the figures be careful as the orientation of the axes is not the same in all instances.

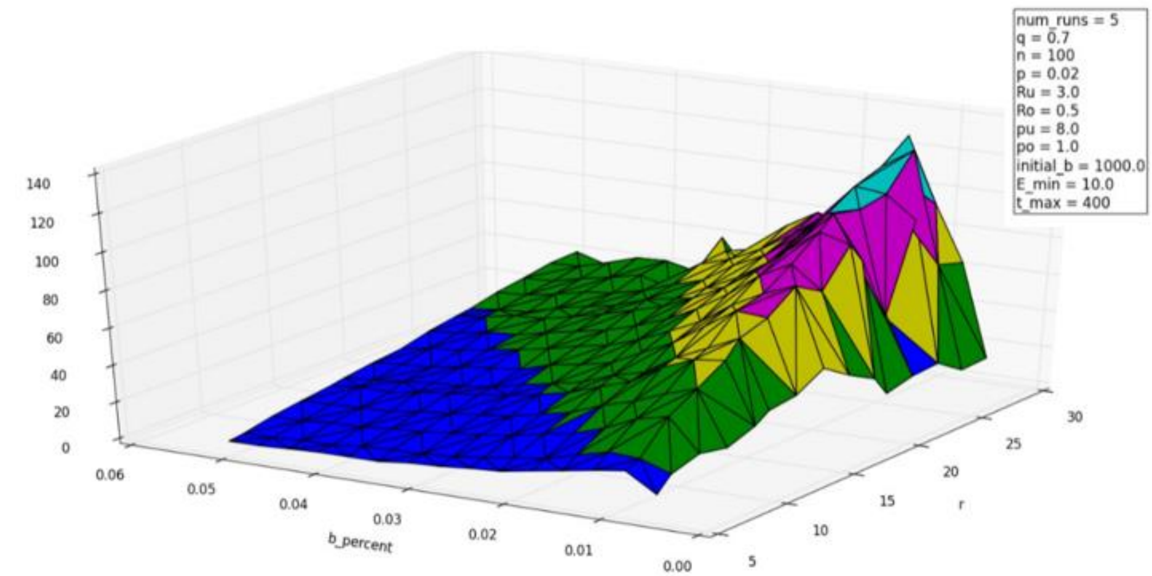


Figure 11: Wealth Metric at the End of the Simulation

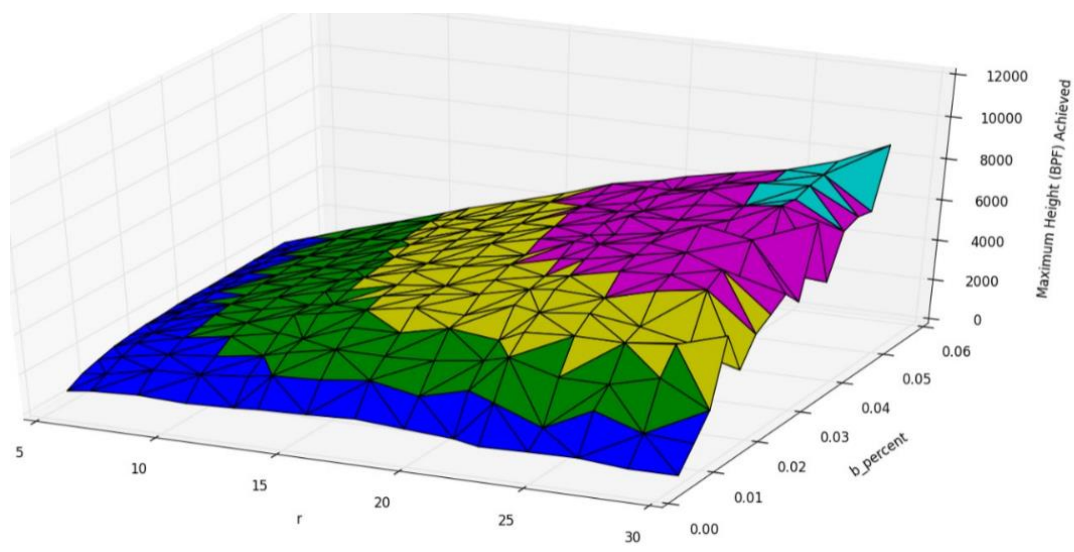


Figure 12: Maximum Height Achieved

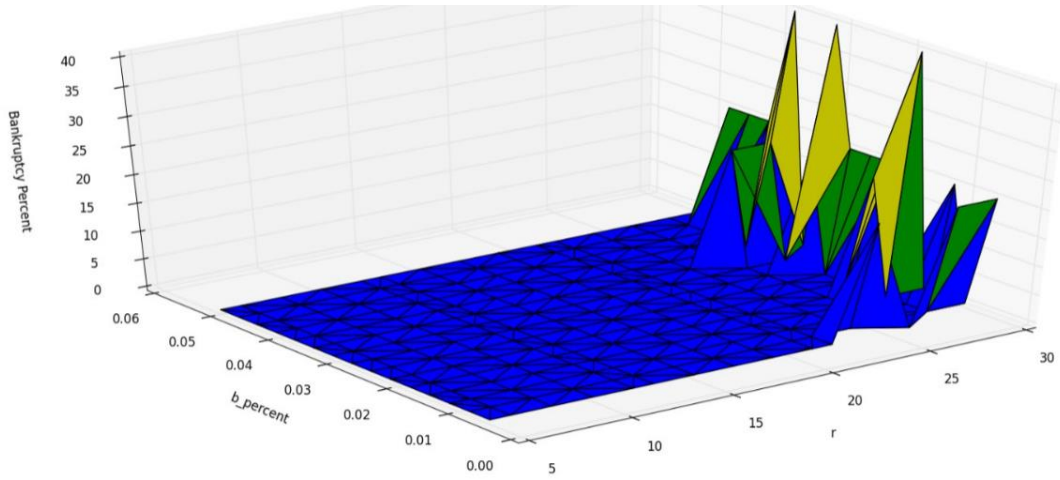


Figure 13: Bankruptcy Percent

To summarize the findings from these figures, let's look at the two parameters individually.

<u>Search Radius</u>

- As r increases both the maximum height and wealth increase
- High values of r may lead the company to go bankrupt

<u>Budget Allocation</u>

- As $b_{\%}$ increases so does the maximum height
- Wealth appears to increase and then decrease with increasing $b_{\%}$
  - Results appear to show that $b_{\%} = 0.1$ maximizes wealth
- Has no apparent effect on bankruptcy

The question arises: How can a company use this information to form a decision strategy that optimizes their performance? There is not a singular answer to this question as the plots demonstrate that some tradeoffs are occurring. The answer, therefore, entirely depends on what the company wishes to achieve. A list of possible objectives and corresponding strategies is listed below in Table 2.

Table 2: Decision Making Strategies for the Company Given an Objective

Objective	Strategy
Maximize Height	High r ; High $b_{\%}$
Maximize Wealth	High r ; Observed Optimal $b_{\%}$
Minimize Bankruptcy & Maximize Height	Moderate r ; High $b_{\%}$
Minimize Bankruptcy & Maximize Wealth	Moderate r ; Observed Optimal $b_{\%}$

The table demonstrates that there are many possible objectives that a company could wish to fulfill. Perhaps a smaller company is risk-averse and therefore wants to minimize their chance for bankruptcy. A company that puts a higher emphasis on creating advanced innovations (maximize height) should follow a different strategy than that of a company focused solely on obtaining the highest return on investment possible (maximize wealth).

4.3 Notable Trends

So far, we have only evaluated performance at a given time point. Below, we will discuss the change of performance metrics over time. We focus on several notable trends that are observed from our simulation results. Correspondingly, the horizontal axis represents time (t), the vertical axis represents the value of the metric being studied, and the different colored lines represent different values parameter being varied. Like in previous sections, multiple simulation runs were conducted for each set of parameter values. In the following instances the results were averaged across ten runs ($num_{runs} = 10$).

4.3.1 Apparent "Takeoff Effect" Exhibited by Wealth

It has been commonly observed in the literature that new product diffusion often features a "takeoff" process.[10] Takeoff, in product diffusion, refers to a new product first experiencing slow sales and adoption by customers. However, after crossing a certain point in time, the sales begin to accelerate and attract more and more customers (i.e. taking off). After this period of rapid growth, the product's sales eventually slow and level off to steady-state.

A similar trend can be seen in our simulation results when looking at the wealth (W_t) growth over time (Figures 14 & 15).

[10] To see percolation models of product diffusion and a more in depth discussion on takeoff see Honisch et al. (2008) and Cantono et al. (2009)

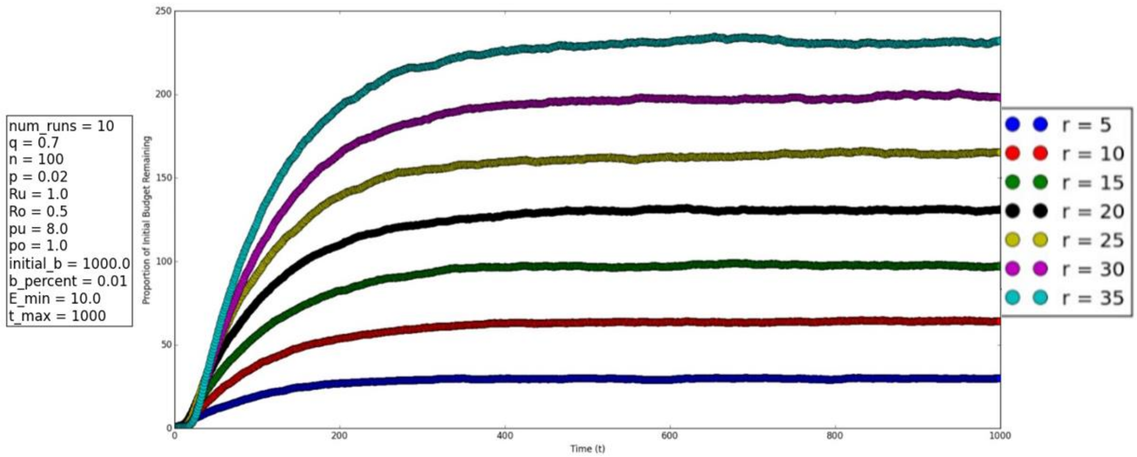


Figure 14: Proportion of Budget Remaining Over Time Varying Search Radius (r)

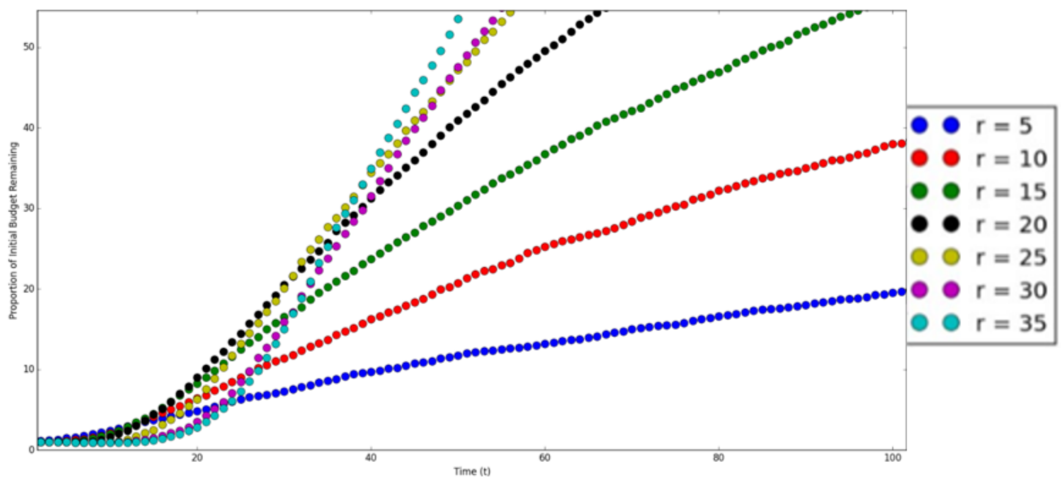


Figure 15:Close-Up of Figure 11 Demonstrating Point When Takeoff Occurs

In all cases the wealth appears to stay the same or decrease in early periods before suddenly increasing rapidly within a short time frame. The rise in wealth eventually slows until reaching a steady-state wealth level. The height that this reaches is dependent on the search radius. Figure 14 demonstrates that if a company chooses a larger search radius they can expect to have a higher steady-state wealth, which means they generate more revenue.

The steady-state of wealth is reached when the returns experienced from finding revenue generators and the cost of performing R&D are equal. These values offset leaving a net gain of zero wealth. The reason that this maximum wealth value increases with increasing r is due to the greater search area that comes along with higher values of r. A larger search radius means more opportunities to find a revenue generator. Also, due to the fact that the effort each period (E) is very large by this point, it overcomes the resistance values of sites easily, even though the effort is distributed over a greater search area.

The search radius also helps determine how long it takes before the growth of wealth enters the acceleration phase. Figure 15 shows that a smaller search radius means a shorter wait before takeoff occurs. A larger search radius means waiting longer for takeoff, but experiencing higher revenues later on. In this regard the company faces a tradeoff between time and money. The question becomes: Do they want to see small revenue generation quickly or larger revenue generation that occurs a little later? The answer to this question depends on many outside factors including the current economic environment, company's specific goals, risk analysis, etc. that go beyond the scope of this paper.

4.3.2 Effect of Company Size on Performance

The width of the lattice structure (n) is representative of how wide of a technical scope a company has to perform research and development search. Therefore, the larger the n value the larger the company. The simulation results discussed in Section 4.1.2 demonstrated that there exists a region of small n values where the likelihood that the company goes bankrupt is high. This section will elaborate more upon this finding by viewing how performance metrics change over time for different values of n.

Figures 16 & 17 demonstrate how the maximum height (i.e. max(BPF)) and wealth change over time for different values of n. These two charts are from the same simulation run with the actual simulation inputs given box seen in Figure 16.

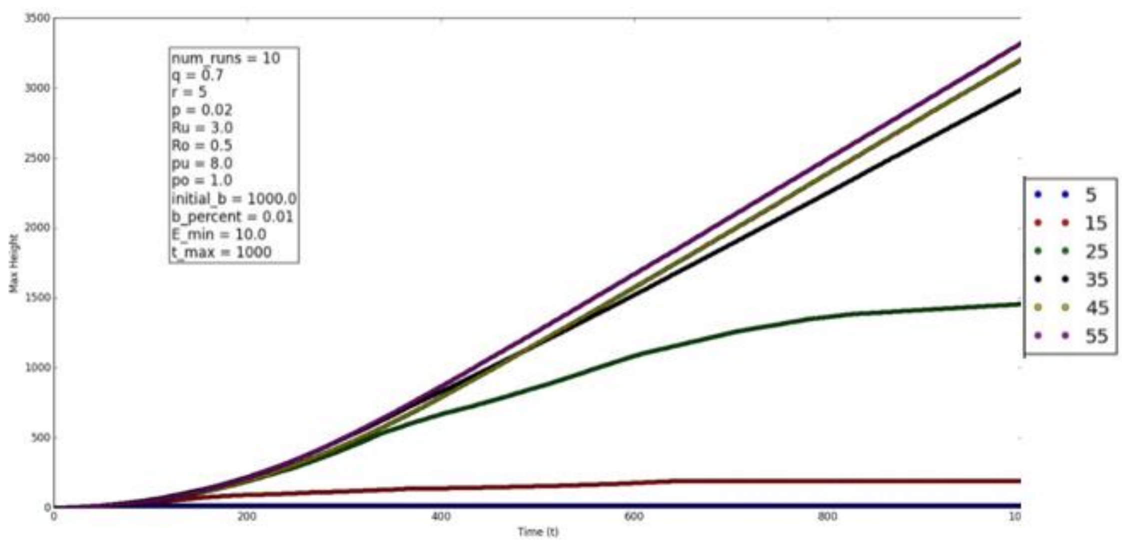


Figure 16: Maximum Height vs. Time for Different n Values

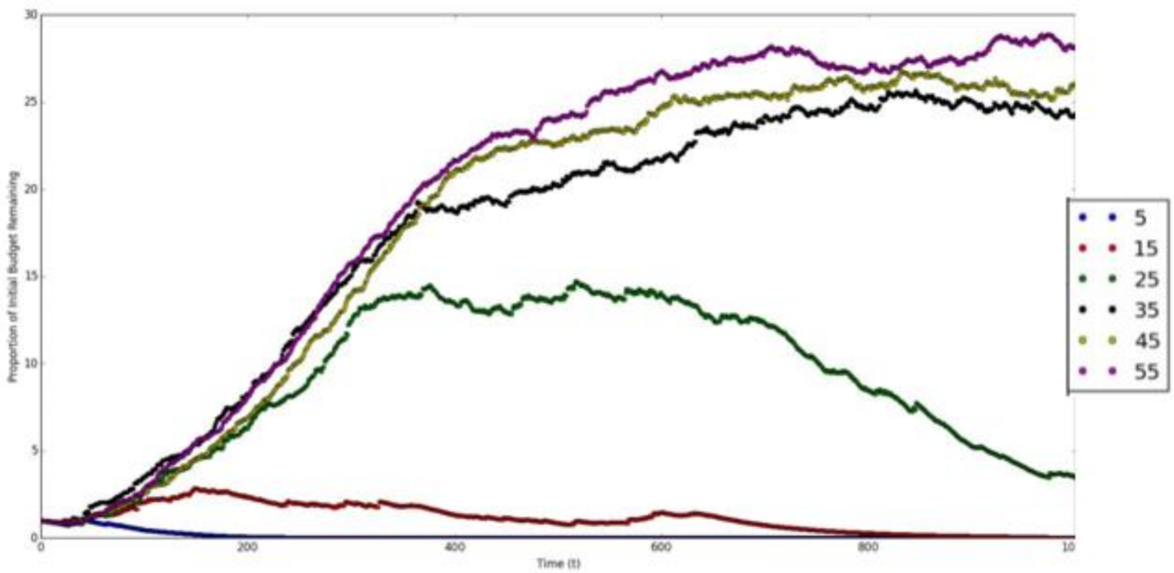


Figure 17: Wealth vs. Time for Different n Values

The first trend to note is that for the two smallest values of n (5 & 15) the company either experiences bankruptcy or comes very close to it. This can be seen by the wealth approaching zero before time period 1000 (t_{max}). This result verifies the previous findings that for small values of n bankruptcy occurs with high probability.

Now that it is apparent that small values of n lead to poor performance we will discuss performance as n increases. The figures demonstrate that both maximum height and wealth increase as n increases. The interesting finding, and the most notable trend from this result, is that this performance gain is ever diminishing with increasing n. To test this finding further, simulations using even larger values of n were conducted using the same input values (Figures 18 & 19). As expected the larger n values provide only marginal improvements in performance. Eventually the performance values converge, so much so in Figure 18 that it is hard to distinguish the different n values from each other.

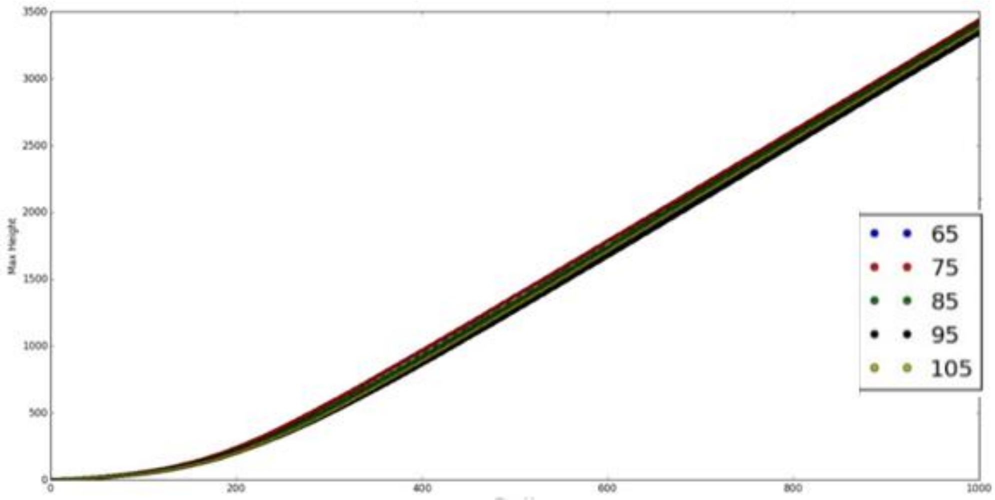


Figure 18: Convergence of Maximum Height for Larger n Values

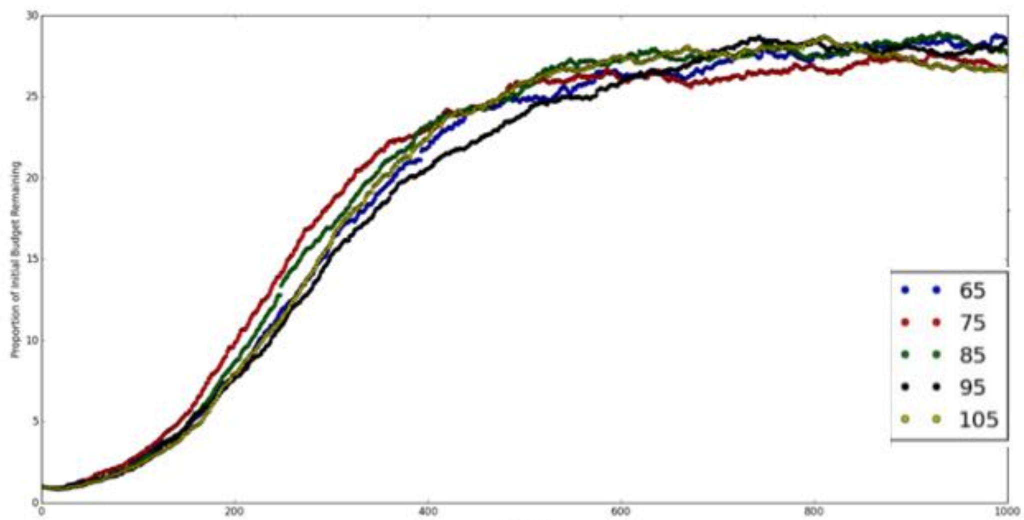


Figure 19: Convergence of Wealth for Larger n Values

This means that a small company with a small total R&D budget (B) that has a narrow technical scope is at a severe disadvantage when compared to a larger company (larger budget & wider scope). However, once a company reaches a certain size (i.e. certain n value), the benefit of being even larger is negligible. In this current model the company cannot actively use this information in decision making because we treat n as a fixed constraint. However, if a company could choose the size of their technical scope at some cost (where a higher n leads to a higher cost) this result could factor into their decision. The company wouldn't necessarily want to pick a very large n, but rather only to the point where the performance appears to converge.

5. Model Variations

This section describes two variations that were made to the model and their subsequent effect on the company's performance. One variation focuses on changing how revenue generators are distributed within the lattice structure. The other changes the structure of budgets and revenue from a company-wide perspective to the individual technological fields (columns) themselves.

5.1 Varying the Distribution of Revenue Generators within the Lattice

In the original model, we have assumed that each site of the lattice has the same probability of being a revenue generator q*p. However, this assumption does not take into account the following "real world" observations:

1. More advanced innovations exhibit higher returns
2. Less advanced innovations are more numerous

To accommodate these changes, the probability (p) that a feasible site is a revenue generator and the revenue values (z) themselves were altered. The variables are now dependent on the row of a given site $i \in [1, \infty)$.

$$p_i = \frac{p}{log_{10}(9+i)} \;\; s.t.\; 0 < p \leq 1 \tag{12}$$

$$z_i = (log_{10}(9+i) * x) \;\; where \;\; x \sim lognormal(p_\mu, p_\sigma) \tag{13}$$

The equations (12 & 13) demonstrate that the probability will decrease and the revenue will increase with higher row values.

Simulation runs were conducted based on the same input values as previously done in Section 4.2. This is so the results can be compared and contrasted between one another to see what effects introduction of row dependence on p & z had on performance. Thus Figures 20 – 22 were compared with the previous Figures 11-13 respectively.

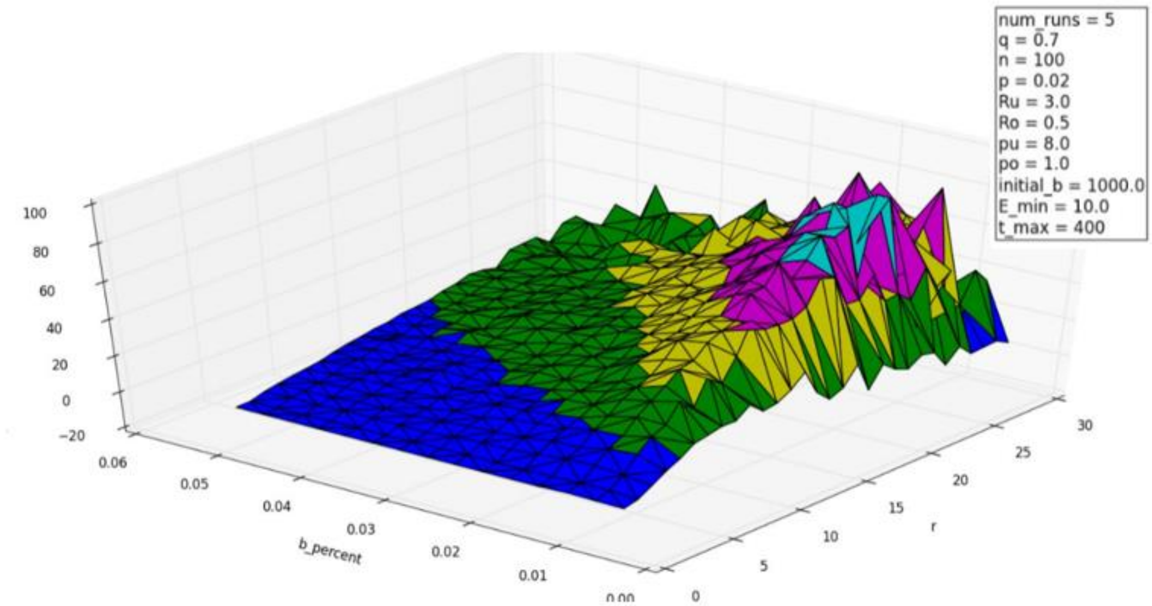


Figure 20: Wealth of Model with Non-Uniform Revenue Generator Placement

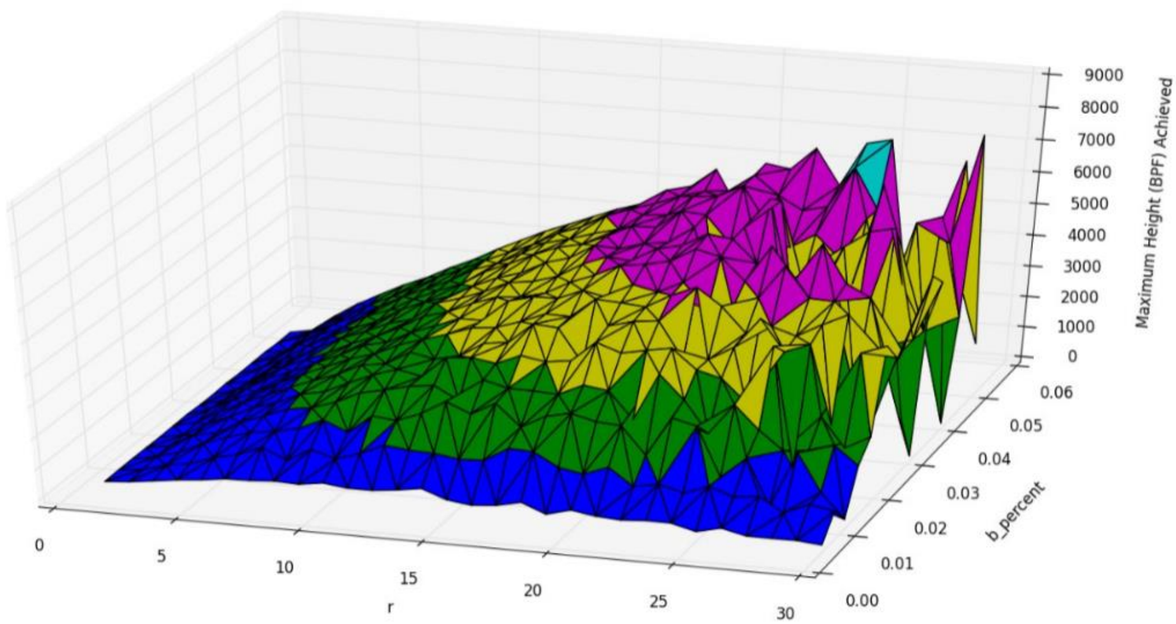


Figure 21: Maximum Height of Model with Non-Uniform Revenue Generator Placement

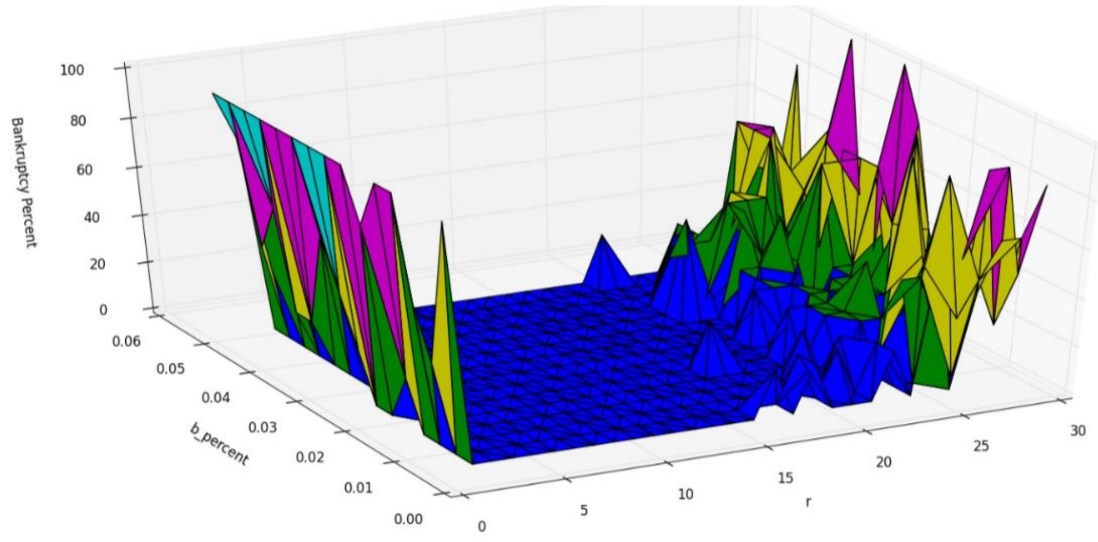


Figure 22: Bankruptcy Percent of Model with Non-Uniform Revenue Generator Placement

Upon initial inspection, the new figures appear to follow the same trends as those previously discussed in Section 4.2.

- **$b_{\%}$ - Percent of Budget Allocated to R&D Search**
  - As $b_{\%}$ increases height increases
  - $b_{\%}$ has an observed optimal value (approximately 0.1 in this example)
- **r - Search Radius**
  - As r increases height (usually) increases
  - As r increases there is an increased chance for bankruptcy to occur

The key difference appears to be that increasing the search radius does not have a universally positive effect on the wealth and maximum height metrics. A possible explanation for this is that a larger r value opens up the possibility of reaching higher rows more quickly. However, with this new variation on the model raising the BPF of a column is not necessarily better because it decreases the chance of finding a revenue generator. On the other hand, operating with lower values of r mean that the company is more likely to stay on lower rows, and even though the expected revenue is lower, the company can more exhaustively search for and expect a higher probability to find sites that generate revenue.

In regards to the possibility of bankruptcy, it appears that larger r values now have a higher chance of causing bankruptcy than in the case of the original model. Figure 22 also demonstrates that now an extremely low r value can also cause bankruptcy as well as high values.

5.2 Separate Budgets per Column

This model variation changes how the R&D budget is portrayed and updated. Previously every time period the total budget (B_t) was allocated equally to each column and any revenue generated that period went back into this total budget. What would happen if instead of a company-wide budget, each column had its own individual budget?

This means that the separate technological fields are no longer "pooling" their budgets together. Also, revenue generators no longer have their capital placed back into the large company budget, but rather the individual budget of the column that discovered it. Note that this model variation still represents a single company, but now only rewards the technical fields that actually discover revenue generators. This is meant to represent a company that actively promotes competition between different R&D teams as an incentive to work harder.

To implement these changes, the equations from Section 2.2 for updating the budget were altered. This alteration allows for each column (j) to have its own corresponding variables:

$$b_{j,0} = \frac{B_0}{n} \ \forall j \ \in [0, n) \tag{14}$$

$$E_{j,t} = \ min(max(b_{j,t} * b_{percent} \ , E_{min}), b_{j,t}) \ \forall j \ \in [0, n) \tag{15}$$

$$b_{j,t+1} = (b_{j,t} - E_t + z_j) \ \ \forall j \ \in [0, n) \tag{16}$$

$$B_t = \sum_{j=0}^{n-1} b_{j,t} \tag{17}$$

If an individual column's budget reaches zero ($b_{j,t}$), before t_{max} periods have been conducted, then that individual column goes "bankrupt" and no longer participates in R&D search. In this regard the Darwinian concept of only the strong survive applies. The idea is that a few strong R&D teams (columns) will emerge and the R&D budget won't go to waste on teams that are underperforming. The company, as a whole, only goes bankrupt if every column has gone bankrupt as well. These conditions can be seen in (18 & 19) respectively.

$$bank_j = 1 \leftrightarrow b_j = 0 \tag{18}$$

$$bank = 1 \ \leftrightarrow \sum_{j=0}^{n-1} bank_j = n \leftrightarrow B_t = 0 \tag{19}$$

The metrics used to gauge the performance of the company are the same as the previous sections. To compare and contrast this variation to the original model, simulations were run using the same parameter values for both the original and new model. Figures 23 & 24 are plots of the company's wealth over time, with varying search radius (r), for the original model & variation respectively. While Figures 25 & 26 are plots from the same simulation run of the maximum height over time, with varying search radius (r), for the original model & variation respectively.

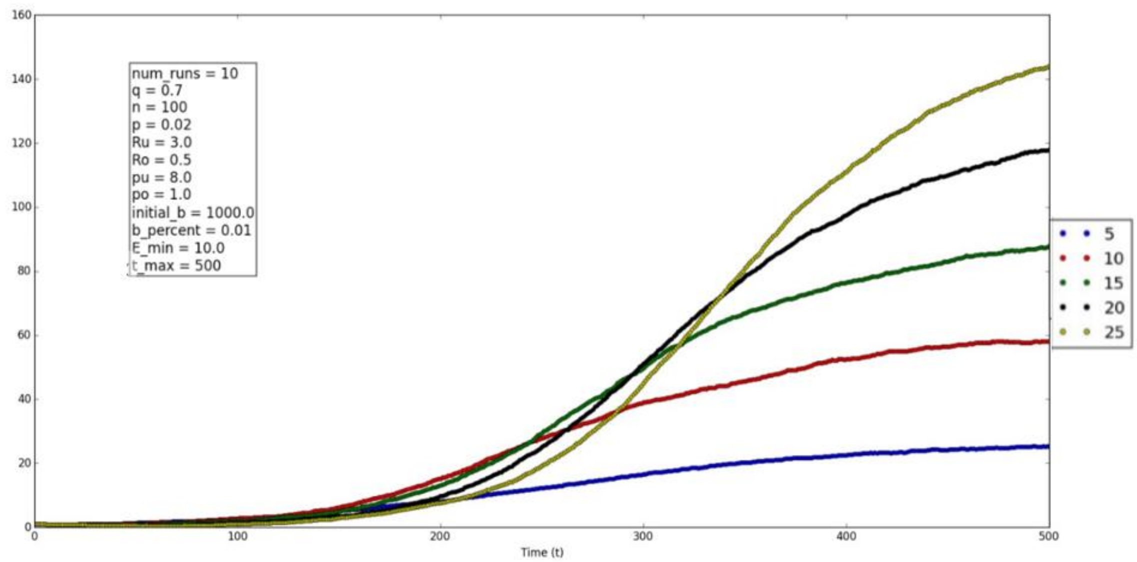


Figure 23: Wealth of Original Model Across Various r Values

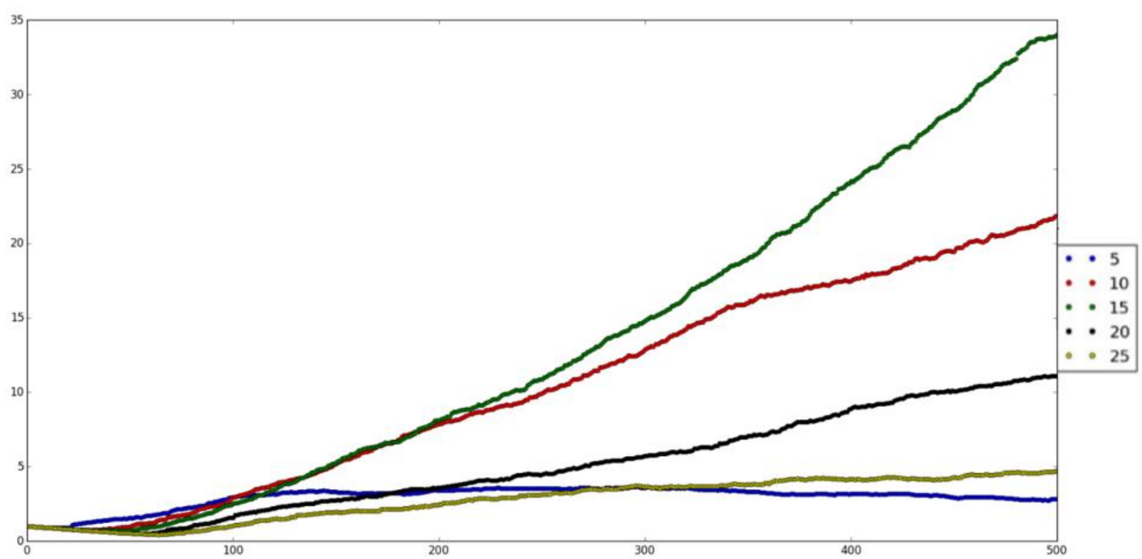


Figure 24: Wealth of Model with Separate Budgets Across Various r Values

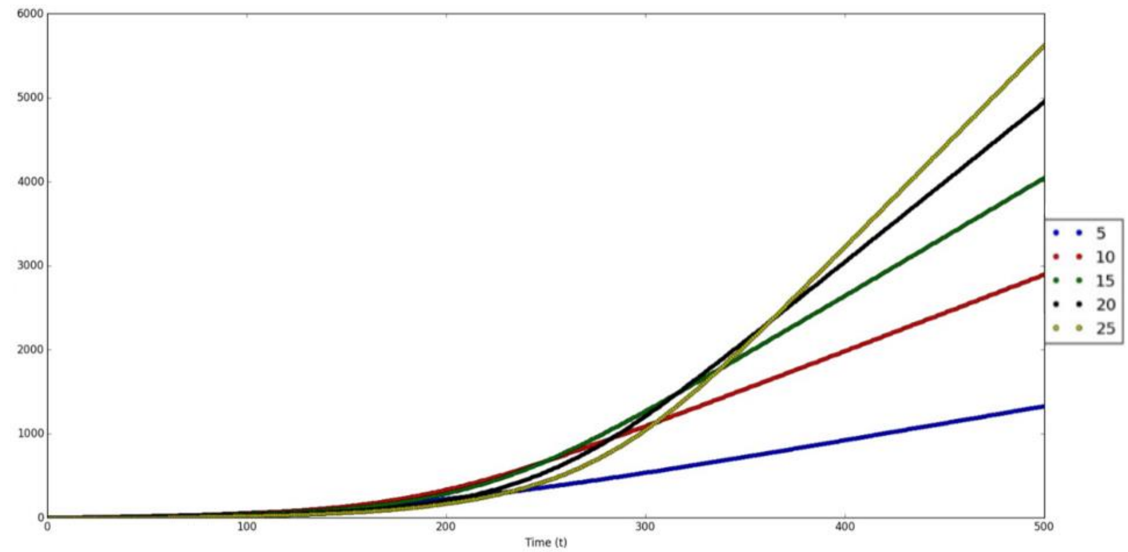


Figure 25: Maximum Height of Original Model Across Various r Values

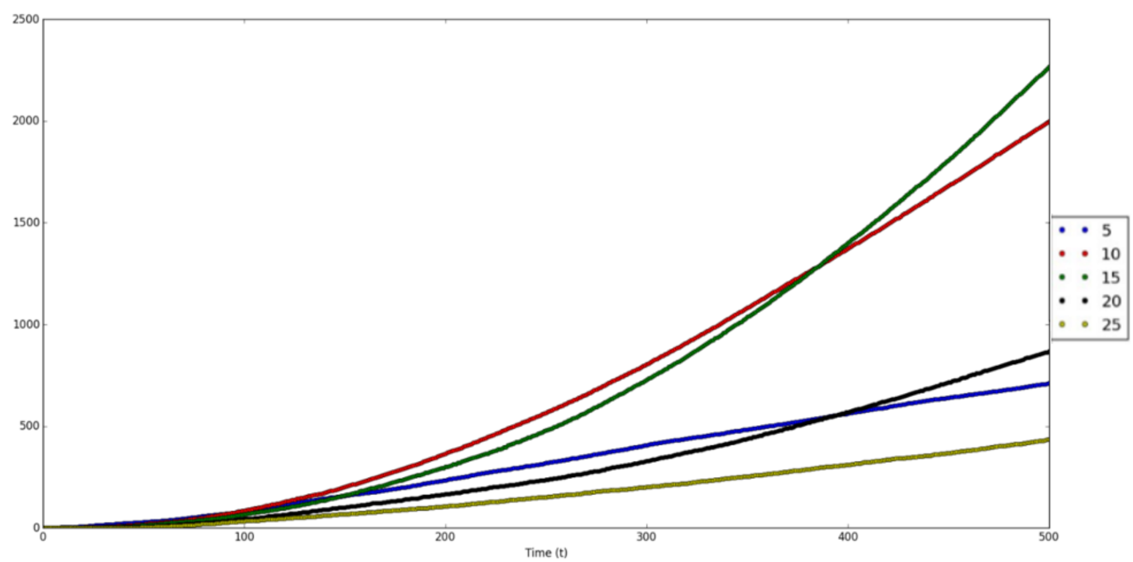


Figure 26: Maximum Height of Model with Separate Budgets Across Various r Values

The figures demonstrate that the new model with separated budgets results in lower values of both performance metrics (wealth & maximum height) across all values of search radius r. As discussed in previous sections, the optimal choice of r for the original is the highest value (r = 25). However, this doesn't appear to hold for the changed model. The optimal r in the separated budget model lies in the middle of the given range (r = 15).

It appears that, once again, a possible explanation for this result is a lack of budget flexibility in the new model when compared to the original. With a budget pool that is shared amongst everyone, specific columns can have multiple periods where they don't find any revenue generators, but still have their budgets supplemented with revenue from their co-workers (i.e. other columns). However, with a competitive work environment that has a "winner takes all" mentality, a column with many consecutive periods lacking revenue generation they will go bankrupt. It would be interesting to find a case where the separate budgets performs better than the original model, but no such cases were found in the many simulations conducted for this paper.

6. Conclusion

This paper carries out a simulation-based study on the company-level innovation process under the framework of a percolation model. Our model is a generalization of the models presented by Silverberg & Verspagen (2003, 2005, 2007). Similar to the latter, we model the innovation process as a search for connected pathways in the technology space (represented by a lattice). We enrich the model by introducing the notion of revenue generators and a wealth evolution process, which lays the foundation for performance analysis.

Our simulation results characterize parameter regions (external limits of survivability) outside of which sustainable wealth growth is attainable for the company. Analysis of this wealth growth over time reveals a trend similar to that of the "takeoff" phenomena. The company's wealth initially declines and then grows rapidly before leveling off to a steady-state value. We discussed different strategies a company can pursue in organizing and investing in the R&D search process. The "best" strategy is heavily dependent on which metric a company wishes to maximize and the model parameters themselves. We demonstrate that smaller companies (less lattice columns) exhibit poorer performance when compared to larger companies, however, this gap in performance is not substantial when comparing medium sized companies to large ones. There appears to be a tradeoff between pushing the technology boundary (i.e. maximum height of the lattice) and exhaustively searching for all possible revenue generators within the technological space. Lastly, companies that attempt to foster competition between R&D groups (i.e. lattice columns) by restricting revenues to only the group that discovered the innovation exhibit poorer performance when compared to companies who take the revenues generated by all groups and pool them together.

Our work opens the door for further investigations. By relaxing assumptions and extending the scope of this model, one may gain many more insights regarding the innovation process. In particular, the following could be considered:

1. Different distributions of the revenue and resistance values
2. Different ways of distributing revenue generators throughout the lattice structure
    2.1 Clustering revenue generators by column
    2.2 Clustering revenue generators in large groups of more than a single site
3. Employ a more sophisticated effort decision (E)
    3.1 Implementation of a maximum effort (E_{max}) the company can put into R&D search
    3.2 Creation of a more complex effort function
4. Allow the company to vary their decisions dynamically each period

5. Further analysis into the benefits/drawbacks of having budgets and revenues for each R&D team (column) rather than the company as a whole. Is there a specific case where this competitive atmosphere leads to better performance?
6. Relax the assumption that effort is equally allocated between all hidden sites in the search area
7. Relax the assumption that total budget is equally allocated between all columns

References

Cantono, Simona, and Gerald Silverberg. "A percolation model of eco-innovation diffusion: the relationship between diffusion, learning economies and subsidies." Technological forecasting and social change 76.4 (2009): 487-496.

Cohendet, P., Zuscovitch, E., 1982. ProgrMes Technique et Percolation. BETA, UniversitXe de Strasbourg Working Paper.

Foray D, Grübler A (1990) Morphological analysis, diffusion and lock-out of technologies: ferrous casting in France and the FRG. Res Policy 19:535–550

Frenken K, Leydesdorff L (2000) Scaling trajectories in civil aircraft 1913–1997. Res Policy 29:331–348

Grimmett, G., 1989. Percolation. Springer, New York.

Harhoff D, Narin F, Scherer FM, Vopel K (1999) Citation frequency and the value of patented inventions. Rev Econ Stat 81:511–515

Harhoff D, Scherer FM, Vopel K (2003) Exploring the tail of patented value distribution. In: Grandstrand O
(ed) Economics, law and intellectual property. Kluwer, Boston/Dordrecht/London

Hohnisch, Martin, Sabine Pittnauer, and Dietrich Stauffer. "A percolation-based model explaining delayed takeoff in new-product diffusion." Industrial and Corporate Change 17.5 (2008): 1001-1017.

Saviotti PP (1996) Technological evolution, variety and the economy. Edward Elgar, Cheltenham and Brookfield

Scherer FM (1998) The size distribution of profits from innovation. Ann Econ Stat 49/50:495–516

Scherer FM, Harhoff D, Kukies J (2000) Uncertainty and the size distribution of rewards from innovation.
J Evol Econ 10:175–200

Schumpeter JA (1939) Business cycles: a theoretical, historical and statistical analysis of the capitalist process. McGraw-Hill, New York (page numbers quoted in the text refer to the abridged version reprinted in 1989 by Porcupine Press, Philadelphia)

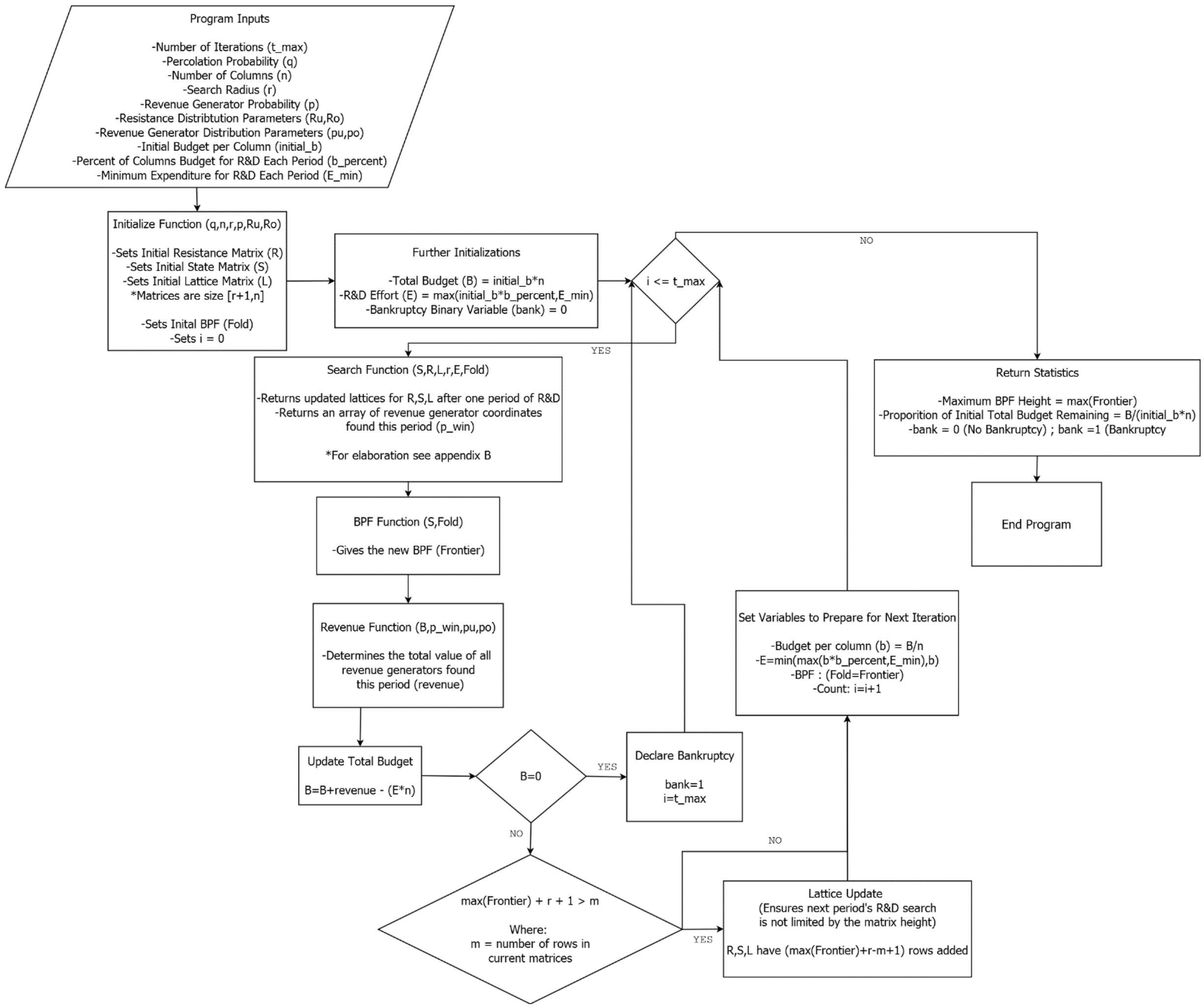
Silverberg, Gerald, and Bart Verspagen. "A percolation model of innovation in complex technology spaces." Journal of Economic Dynamics and Control29.1 (2005): 225-244.

Silverberg, Gerald, and Bart Verspagen. "Brewing the future: stylized facts about innovation and their confrontation with a percolation model." EMAEE Conference, Augsburg. 2003.

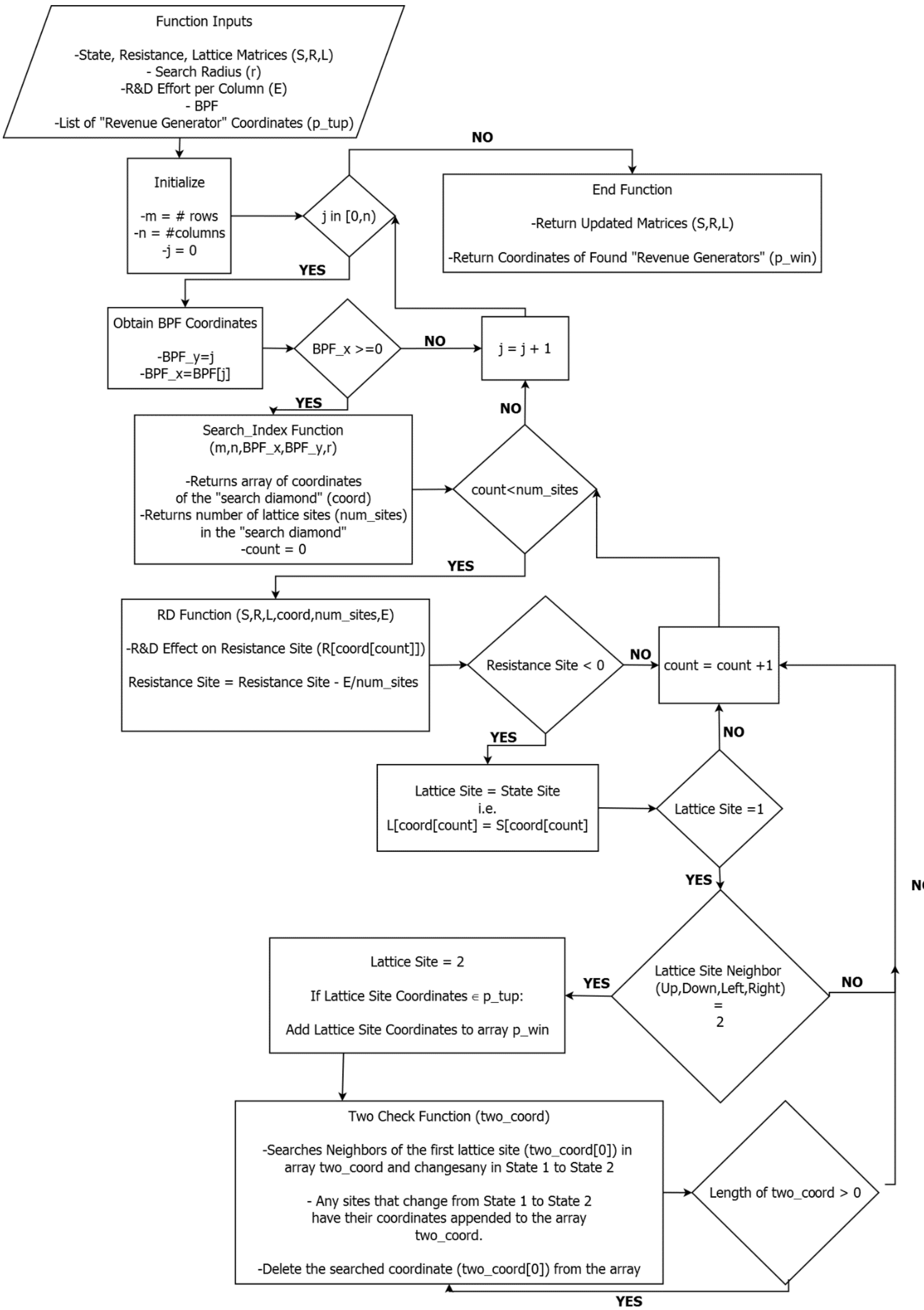
Silverberg, Gerald, and Bart Verspagen. "Self-organization of R&D search in complex technology spaces." Journal of economic interaction and coordination 2.2 (2007): 211-229.

Trajtenberg M (1990) A penny for your quotes: patent citations and the value of innovations. Rand J Econ 21:172–187

Appendix A: Algorithm Flow Chart

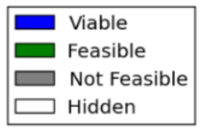


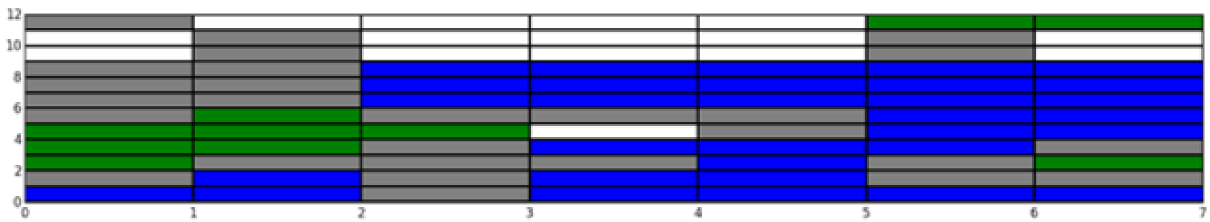
Appendix B: Search Function Flow Chart



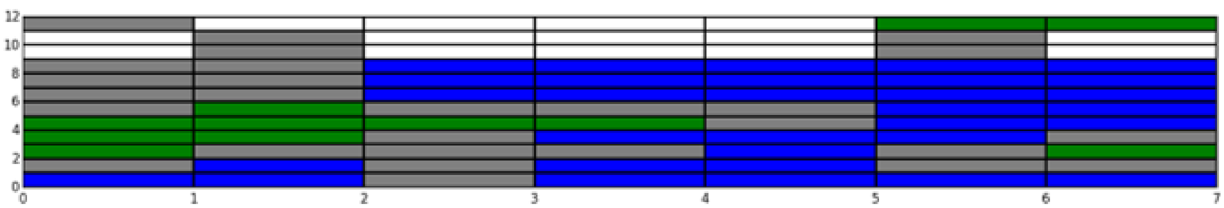
Appendix C: Specific R&D Search Example

The lattices below are snap shots of what the lattice looks like as it is being updated during the R&D Search step of the algorithm. Focus your attention to the hidden (white) site in the middle of the first lattice ($S_{3,4} = -1$). This will be the focus of the example and represents how a "cornerstone innovation" can cause multiple innovations to appear at once.

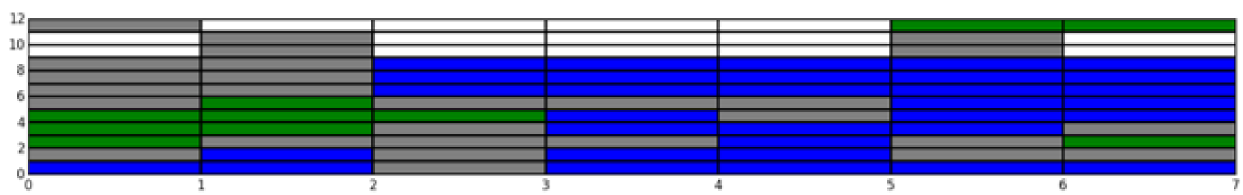




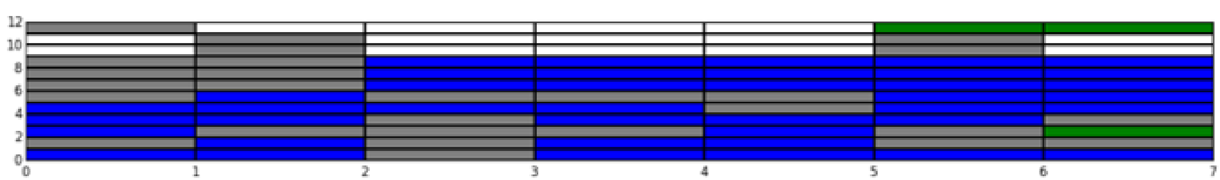
R&D Search Begins



Resistance Value Becomes Negative Changing Site from State -1 to State 1



Checking Neighbors Finds Bottom Site Already in State 2 so Site Changes from State 1 to State 2



Two Check Function First Changes the Left Neighbor to State 2 and then Loops Creating Many Innovations (See Left Side of Matrix)